



Étudier la géographie des activités et des collectifs scientifiques dans le monde : de la croissance du système de production contemporain aux dynamiques d'une spécialité, la réparation de l'ADN

Marion Maisonobe

► To cite this version:

Marion Maisonobe. Étudier la géographie des activités et des collectifs scientifiques dans le monde : de la croissance du système de production contemporain aux dynamiques d'une spécialité, la réparation de l'ADN. Géographie. Université Toulouse le Mirail - Toulouse II, 2015. Français. NNT : 2015TOU20055 . tel-01235015

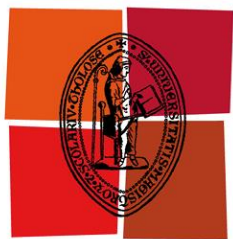
HAL Id: tel-01235015

<https://theses.hal.science/tel-01235015>

Submitted on 27 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse Jean Jaurès

Présentée et soutenue par

Marion MAISONOBE

Le 17 septembre 2015

Étudier la géographie des activités et des collectifs scientifiques dans
le monde

*De la croissance du système de production contemporain aux dynamiques
d'une spécialité : la réparation de l'ADN*

École doctorale et discipline ou spécialité

ED TESC – Géographie - Aménagement

Unité de recherche

LISST – CIEU – UMR 5193 CNRS – EHESS – Université Toulouse Jean Jaurès

Directeur de Thèse

ECKERT Denis – Directeur de recherche – CNRS – LISST CIEU – UMR CNRS 5193

Jury

BOUBA-OLGA Olivier – Professeur des Universités – Université de Poitiers – **Rapporteur**

GRASLAND Claude – Professeur des Universités – Université Paris 7 Denis Diderot – **Rapporteur**

BARON Myriam – Professeur des Universités – Université Paris Est Créteil – **Membre**

EVENO Emmanuel – Professeur des Universités – Université de Toulouse Jean-Jaurès – **Membre**

GINGRAS Yves – Professeur des Universités – Université du Québec à Montréal – **Membre**

MILARD Béatrice – Professeur des Universités – Université de Toulouse Jean-Jaurès – **Membre**

Université Toulouse Jean Jaurès - Le Mirail

Laboratoire Interdisciplinaire Solidarités, Sociétés, Territoires

Centre Interdisciplinaire d'Études Urbaines

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ

EN GÉOGRAPHIE – AMÉNAGEMENT

Étudier la géographie des activités et des collectifs scientifiques dans le monde.

De la croissance du système de production scientifique contemporain

Aux dynamiques d'une spécialité : la réparation de l'ADN

Maisonobe Marion

Présentée et soutenue publiquement

Le 17 septembre 2015

Directeur de Recherche

ECKERT Denis, Directeur de recherche au CNRS

JURY

BOUBA-OLGA Olivier – Professeur des Universités – Université de Poitiers – **Rapporteur**

GRASLAND Claude – Professeur des Universités – Université Paris 7 Denis Diderot – **Rapporteur**

BARON Myriam – Professeur des Universités – Université Paris Est Créteil – **Membre**

EVENO Emmanuel – Professeur des Universités – Université Toulouse Jean Jaurès – **Membre**

GINGRAS Yves – Professeur des Universités – Université du Québec à Montréal – **Membre**

MILARD Béatrice – Professeur des Universités – Université Toulouse Jean Jaurès – **Membre**

« Ideas which are not cultivated by people whose regular jobs are to cultivate them are like souls hovering in a mythological limbo before entering a body. They can light upon the dreams or the imagination of one person here and another one there, of someone who lives today or of someone else who will be born in a thousand years. If however, ideas become the end-products of scientific roles, they can be likened to genes which are transmitted from generation to generation through a reliable and natural process; under normal conditions, they will not only survive but increases. »

Ben-David & Collins, 1966

Remerciements

Avant toute chose, merci aux deux Laurents : Beauguitte et Jégou, pour leur aide et leur influence décisive tout au long de ce travail. L'histoire débute entre Paris et Toulouse en même temps que le programme de recherche « Géoscience ». Le financement de thèse s'intégrant dans la dynamique de ce programme, me voilà plongée au milieu d'une marée de spécialistes des sciences aux diverses origines géographiques et disciplines scientifiques, à qui cette thèse doit énormément. Parmi les nombreux participants qui l'ont influencée, je voudrais remercier Denis Eckert et Michel Grossetti ayant participé à la coordination du programme de recherche et à celle de ma thèse (surtout pour le 1^{er} !) avec un enthousiasme sans faille, et si essentiel pour arriver jusqu'au bout ! Myriam Baron et Béatrice Milard pour leurs précieux conseils, et comme collaboratrices. Merci, aussi, à André Grelon pour les passionnantes réunions de la tâche 1, à René Sigrist, Natalia Tikhonov, Pierre Mounier-Kuhn, Caroline Barrera, Yves Lequin, Mina Kleiche, Roland Waast, Marie-Noëlle Pane, Dominique Vinck, Vincent Simoulin, Muriel Lefebvre, Guillaume Cabanac, Corinne Bonnet pour les discussions. L'OST Montréal et Paris, pour les données !

Les premières années ont été mouvementées, et je remercie l'UMS RIATE de m'avoir prêté un bureau dans leurs conviviaux locaux de Tolbiac. Merci à Bernard Corminboeuf, pour sa confiance et l'opportunité qu'il m'a donnée de travailler pour la revue Netcom. Entre Paris et Toulouse, il y a le laboratoire des sciences de l'innovation à Eindhoven (ECIS), et Koen Frenken, son directeur d'alors, à qui cette thèse doit beaucoup. Là-bas, j'aimerais remercier J. Hoekman, S. Hardeman et F. Veraart pour leur aide, tout le collectif « jeune » pour les déjeuners, avec une mention spéciale pour Rudi Bekkers et Yuti Ariani pour les discussions infinies ! Enfin, je remercie les spécialistes néerlandais de la réparation de l'ADN qui m'ont raconté leurs trajectoires. Ensuite, croyez-le ou non, il y a Athènes, Montréal, Rabat, Porquerolles, Dourdan, Utrecht et Ferney-Voltaire : ce n'étaient que des colloques, mais je remercie sincèrement toutes les personnalités croisées là-bas, les discussions, les présentations : j'ai essayé de ne pas en rater une goutte ! Et pendant ce temps, il y a le groupe fmr, vous me suivez ? Non, ce n'est pas grave : merci Laurent, Matthieu, César, Françoise et Serge. Je vous dois tellement. Enfin, il y a Toulouse, le LISST, le CIEU, l'atelier Carto, SMS, le groupe ResTo, Ambra et l'ADN, Pauline et les satellites, Sébastien et le groupe pain, Dorothée et la gentiane, Jules ou la musique, Tristan, la philosophie et le marxisme, Henri et les plantes vertes, Adrien, Julie et la Zumba, Anne, Amira, David et le monde arabe, Josselin, Julie, Ina, Jéro, Constanza, Samuel et Delphine, Henri j'insiste, avec Delio et l'InternaZionale, Edmond et TESC, les doctorants CERS, Marie-Ange Parisot, l'ange Akermann, Annalisa, Léo, merci, merci !

Venons-en à Nice : mes chers parents, je vous dois tant ! Mamie, Pierrette, Alexandra et Nico, les Mi, Mireille... je n'ai jamais assez de temps à vous consacrer ; à Bulle, mon parrain et sa petite famille ; Agay ; Fort-Mahon plage ; le Pouliguen et ses pingouins ; le phoque de Lacanau ; retournons à Paris : mes beaux neveux, Poilou et Aurélie, Rémi, les copains d'aménagement, les cousins, vous êtes magiques ! Passons un instant par Lyon ou Paris, je ne sais plus : Doudou, copine Julie, Maria, vous m'êtes indispensables ! Dépassons tous les clivages : les Marions (la blonde dite Mayon et la rousse qui voyage beaucoup), Benoît et Mickaël : où en serais-je sans vous ? Et revenons à Toulouse, à moins que ce ne soit la Réunion ou l'Afrique : Fabrice Pousse, tu es un cap, que dis-je ! ? Pierre et Pauline, merci pour votre accueil en Occitanie ; la collocation à Foun Lestang ! Merci aussi : Yéyé et le Saint-Louis, Kewin et le cinéma, Marie-Pierre et le dessin, Tom, la poésie et le théâtre... bref la rescousse ! Terminons par la fin, une mention spéciale aux passagers survivants des dernières heures, ceux qui savent, Simon Labussière, le relecteur masqué ! Puis tout là-haut à mes grands-pères, Bridget, Pratchett, J.R.R. Tolkien, Lou Reed...

TABLE DES MATIÈRES

INTRODUCTION	1
PARTIE 1. PENSER LA SCIENCE À L'ÉCHELLE DU MONDE EN GÉOGRAPHIE ...	10
CHAPITRE 1. SCIENCES DE LA SCIENCE, SCIENCES SOCIALES ET GÉOGRAPHIE	9
1. La science comme fait social	10
1.1. <i>Social versus cognitif ?</i>	10
1.2. <i>Changements de « programmes »</i>	12
2. La géographie comme science humaine et sociale	16
2.1. <i>Le « tournant culturel » en géographie</i>	16
2.2. <i>Position sociale, position spatiale et réflexivité</i>	23
CHAPITRE 2. PETITE HISTOIRE DE LA SCIENCE EN GÉOGRAPHIE.....	27
1. Idéologie dominante et obstacles pour aborder la science	28
1.1. <i>La mesure de la science en géographie</i>	28
1.2. <i>L'approche de l'OCDE : enjeux économiques et politiques</i>	32
1.3. <i>La science : ressource ou richesse ?</i>	36
2. De la diffusion des innovations à la géographie de l'innovation	39
2.1. <i>De la diffusion des innovations</i>	39
2.2. <i>...à la géographie de l'innovation</i>	44
CHAPITRE 3. VERS UNE GÉOGRAPHIE DES SCIENCES.....	51
1. La géographie de la recherche : entre logiques d'innovation et de formation	52
1.1. <i>Au-delà d'une géographie de l'innovation...</i>	52
1.2. <i>...et d'une géographie de l'enseignement supérieur</i>	58
2. La géographie de la recherche : quelle est la place des logiques scientifiques ?	66
2.1. <i>Pôles scientifiques et gouvernance universitaire</i>	66
2.2. <i>Villes, réseaux de villes et activités scientifiques</i>	72
CHAPITRE 4. ÉTAT DES SAVOIRS POUR UNE GÉOGRAPHIE DES SCIENCES	83
1. La science dans l'espace et dans le temps avant les années 1990	84
1.1. <i>De la communauté scientifique</i>	84
1.2. <i>...aux communautés scientifiques</i>	90
2. Les sciences dans l'espace et dans le temps à partir des années 1990.....	95
2.1. <i>Du « tournant spatial » en science studies</i>	95
2.2. <i>...à la « scientométrie spatiale »</i>	102

CHAPITRE 5. CADRE ET SUPPORT DU TRAVAIL : LE PROGRAMME « GÉOSCIENCE » ET LE WEB OF SCIENCE..... 109

1. Le programme « Géoscience » 110
 - 1.1. Une approche mixte et interdisciplinaire 110
 - 1.2. Une analyse quantitative de l'évolution de la répartition spatiale des publications scientifiques et des collaborations (tâche 2) 119
2. Une source tertiaire: le *Web of Science* 129
 - 2.1. « L'ère romantique » de l'*Institute for Scientific Information (ISI)* ou les origines du *Web of Science (WoS)*..... 129
 - 2.2. Les données de l'*ISI* : reflet de la science mondiale ? 138

PARTIE 2. LA GÉOGRAPHIE DES SCIENCES À TRAVERS LE WEB OF SCIENCE.157

CHAPITRE 6. LOCALISER LES DONNÉES BIBLIOGRAPHIQUES 159

1. Les lieux de la production scientifique 160
 - 1.1. Les traitements par organisation 160
 - 1.2. Les traitements par entité spatiale 165
 - 1.3. Le géocodage des données bibliographiques 170
2. La délimitation des agglomérations scientifiques 177
 - 2.1. Définir la ville et délimiter des zones urbaines 179
 - 2.2. Une méthode adaptée à l'étude des sciences 190

CHAPITRE 7. SYNTHÉTISER LES DONNÉES BIBLIOGRAPHIQUES 201

1. Le traitement des données de production 202
 - 1.1. Le comptage de la production scientifique..... 202
 - 1.2. L'exemple de la production néerlandaise 206
 - 1.3. Les politiques de normalisation des adresses..... 212
2. Le traitement des données de collaboration 217
 - 2.1. Du réseau de co-signatures au réseau de lieux 217
 - 2.2. Agrégation spatiale versus cohésion scientifique..... 223

CHAPITRE 8. L'ANALYSE SPATIALE DES DONNÉES BIBLIOGRAPHIQUES..... 233

1. La visualisation et la représentation des données : 234
 - 1.1. La visualisation de « big data » à l'échelle mondiale 234
 - 1.2. La représentation comme résultat de recherche 242
2. Les méthodes d'analyse spatiale quantitative et de réseaux 253
 - 2.1. La répartition spatiale des activités scientifiques..... 256
 - 2.2. L'organisation spatiale des activités scientifiques 269
 - 2.3. Expliquer la géographie des activités scientifiques 283

CHAPITRE 9. LES ÉVOLUTIONS DE LA GÉOGRAPHIE DES ACTIVITÉS SCIENTIFIQUES	
CONTEMPORAINES	295
1. Géographie de la production scientifique contemporaine	296
1.1. <i>Une approche multi-niveaux pour saisir la croissance de la science</i>	296
1.2. <i>La déconcentration des activités scientifiques à plusieurs niveaux</i>	301
1.3. <i>Nationalisation et internationalisation des activités scientifiques</i>	309
2. Géographie du système mondial des collaborations scientifiques	319
2.1. <i>La densification du réseau de villes</i>	319
2.2. <i>Des aires géographiques de collaborations privilégiées</i>	329
PARTIE 3. UNE SPÉCIALITÉ SCIENTIFIQUE, « LA RÉPARATION DE L'ADN »...343	
CHAPITRE 10. L'ÉMERGENCE ET LA DIFFUSION D'UNE SPÉCIALITÉ SCIENTIFIQUE : LA RÉPARATION DE L'ADN	345
1. L'étude d'une spécialité scientifique	346
1.1. <i>Défis et limites de l'approche bibliographique</i>	346
1.2. <i>La géographie de la réparation de l'ADN : quelle évolution ?</i>	353
1.3. <i>La géographie de la réparation de l'ADN : quel poids du passé ?</i>	358
2. L'émergence de la spécialité « DNA Repair »	368
2.1. <i>Les trajectoires des pionniers de la spécialité</i>	368
2.2. <i>La rencontre entre le mythe et les premières traces d'activité</i>	376
2.3. <i>La « communauté scientifique » comme un réseau de lieux</i>	386
CHAPITRE 11. LA DIFFUSION SPATIALE D'UNE QUESTION DE RECHERCHE	397
1. Géographie de l'étude du lien entre la transcription et la réparation de l'ADN	398
1.1. <i>L'expansion géographique de la question de recherche</i>	398
1.2. <i>L'organisation spatiale de la question de recherche</i>	405
2. Du point de vue d'une chercheuse toulousaine	414
2.1. <i>« Faire partie de la légende », puis revenir en France</i>	414
2.2. <i>Une publication co-écrite entre Rotterdam et Toulouse</i>	418
2.3. <i>La géographie du réseau et de l'univers de référence de la publication</i>	422
CONCLUSION	431
BIBLIOGRAPHIE	444
LISTE DES PERSONNES RENCONTRÉES	489
LISTE DES TABLEAUX	494
TABLE DES FIGURES.....	496

INTRODUCTION

Les laboratoires de recherche et les universités se sont répandus au XX^e siècle. Leur nombre et leur diversité augmentant, c'est pour assurer une organisation du travail optimale que les dispositifs de coordination se sont multipliés à tous les niveaux institutionnels. Parmi les raisons internes de ce besoin d'organisation, il y a la crainte de faire des découvertes redondantes et, pour les scientifiques, de passer à côté de travaux proches de leurs préoccupations. À mesure que les moyens technologiques de communication et d'accès à l'information se sont diffusés, certains se sont mis à parler de « collège global » et le romancier David Lodge est l'un des premiers à évoquer, dès les années 1980 (soit avant l'avènement d'Internet) les dessous d'un « campus global » (Lodge, 1984). Il faut dire qu'avec la « croissance de la science », expression permettant de qualifier l'augmentation continue dans plusieurs pays, depuis le XIX^e siècle, du nombre de chercheurs et de leurs publications, l'organisation des scientifiques est devenue l'objet de recherches scientifiques (Ben-David, 1991). Les spécialistes des études sur la science issus de la sociologie, de l'histoire ou encore de la linguistique se sont retrouvés au sein d'un champ scientifique hybride connu dans le monde anglo-saxon sous le nom de *science studies*. Ce domaine de recherche vieux d'une cinquantaine d'années a permis qu'ait lieu une forme de démystification de l'activité scientifique. Il est devenu possible en sciences sociales, en dépit du travail de réflexivité que cette démarche suppose, d'aborder l'activité scientifique comme un secteur d'activité, un monde social, une communauté de pratiques comme les autres.

Dans la mouvance du « tournant culturel », les études des sciences se sont orientées vers une approche spatiale à partir des années 1990 (Livingstone, 2003). Il existe plusieurs raisons qui justifient l'intérêt croissant pour une géographie des activités scientifiques : le besoin politique de rationaliser l'effort financier en faveur d'un nombre toujours plus important de sites et d'acteurs scientifiques, la nécessité interne à la science de situer de potentiels collaborateurs ou compétiteurs avec lesquels interagir ou s'organiser pour faire cause commune et, enfin, l'intérêt pour les sciences sociales de mieux saisir l'inscription spatiale des activités scientifiques. Pourtant, comme le remarquaient récemment Jean-Marc Besse à propos de la géographie culturelle d'une part, Denis Eckert et Myriam Baron à propos de la géographie quantitative d'autre part, les géographes français ne se sont pas encore saisis de la question des sciences (Besse, 2010 ; Eckert & Baron, 2013). Les travaux en français de Xavier Polanco sur la « science-monde », de Michel Zitt ou d'Yves Gingras sur l'« internationalisation » de la science ont été peu commentés, pour ne pas dire ignorés, par la géographie (Polanco, 1990 ; Zitt & Bassecoulard, 1999 ; Gingras, 2002). Cette inattention est d'autant moins justifiée qu'à plusieurs reprises dans ses travaux sur la mondialisation, Olivier Dollfus soulignait

l'intérêt d'une géographie des sciences : « Par ses structures et ses rivalités, ses enjeux aussi, sa répartition géographique, la "communauté scientifique" est un intéressant microcosme des réalités de la mondialisation contemporaine. » (Dollfus, 1997, p. 119).

Cette thèse entend prendre la mesure de ce qui a été fait dans les autres disciplines pour s'attacher à montrer qu'il y a en géographie des savoirs et des méthodes susceptibles de nourrir et d'enrichir une réflexion sur l'état et les dynamiques des systèmes scientifiques à l'échelle mondiale. Il conviendra de prouver que la géographie a une légitimité pour produire un discours sur l'inscription spatiale des activités scientifiques, qu'elle offre un cadre tout à fait adapté pour comprendre l'articulation des différents niveaux spatiaux d'organisation de l'activité scientifique, qu'elle est parfaitement armée pour questionner les discours sur la mondialisation du « champ scientifique » (l'expression est de Pierre Bourdieu). Nous irons chercher dans la littérature existante ce qui peut être repris, ce qui peut être enrichi mais aussi ce qui doit être rejeté par notre approche géographique. Une fois que ce travail de fouille sera achevé et qu'il sera admis qu'une géographie des sciences a sa place d'une part en géographie et d'autre part dans le cadre des *science studies*, nous nous emploierons à justifier et proposer une démarche adaptée à l'étude de la répartition des activités scientifiques dans le monde et de la structuration dans l'espace géographique de collectifs scientifiques. L'approche que nous avons retenue pour cette thèse, qui n'est qu'une approche possible parmi l'éventail de possibilités ouvert par le développement d'une branche de géographie dédiée à l'étude des rapports entre sciences et espace géographique, est fondée sur trois propositions.

La première de ces propositions est d'appréhender l'activité scientifique par le biais des productions paraissant dans les revues scientifiques, ce qui d'emblée donne à cette thèse une dimension quantitative et empirique. Ainsi, on se fixe comme objectif de localiser et d'analyser l'organisation spatiale des activités scientifiques par le traitement et l'exploitation de notices bibliographiques accessibles dans les grandes bases privées ayant vocation à couvrir le fleuron des publications scientifiques mondiales (livres exceptés) : le *Web of Science*, propriété de Thomson Reuters, et *Scopus*, géré par l'éditeur scientifique Elsevier. La plus connue, le *Web of Science*, qui est la plus utilisée pour l'établissement des classements internationaux, fera l'objet d'une attention toute particulière. C'est aussi celle sur laquelle notre équipe de recherche s'est appuyée pour construire une population d'unités élémentaires d'analyse baptisées « agglomérations scientifiques ». Nous aurons à cœur de montrer que cette base de données ne reflète qu'une partie seulement de l'activité scientifique qui, du fait de la médiatisation dont elle bénéficie, peut être qualifiée de « science visible ». L'élargissement progressif de la couverture géographique de la base sera envisagé comme un objet d'étude en analysant les contraintes que pose la variabilité de son périmètre. En effet, la couverture de la base rend compte des stratégies de visibilité

prises en œuvre par les comités éditoriaux de revues qui doivent formuler une demande pour figurer dans la base. En plus, compte tenu de l'incitation pesant sur les chercheurs de certains pays à publier dans les revues « visibles », leur progression dans la base n'est pas indépendante d'une volonté étatique de peser dans ce que d'aucuns qualifient de « science mondiale ». Considérer la base de données comme un objet qui est, en soi, digne d'intérêt pour le géographe curieux de comprendre et d'interpréter les dynamiques mondiales constitue l'une des trois propositions qui structurent cette thèse.

La deuxième proposition consiste à prendre des entités urbaines comparables comme unité spatiale élémentaire pour l'analyse. Pour des analyses au niveau mondial, il est nécessaire de réfléchir à un niveau de résolution qui soit le même quel que soit le pays considéré et, par conséquent, qui s'affranchisse des découpages administratifs dont les critères sont propres à chaque pays. De nombreux travaux de géographes (Cattan & Rozenblat, 1991 ; Moriconi-Ebrard, 1991) attirent l'attention sur le fait que l'hétérogénéité des découpages administratifs rend délicate toute comparaison internationale des villes. Pour autant, les classements internationaux de villes prolifèrent qui ignorent ce problème et prétendent informer la hiérarchie des « villes mondiales ». Prendre au sérieux la question du niveau d'agrégation des données est primordial dans un contexte où le couple global-local nourrit nombre de travaux scientifiques (Sassen, 1991 ; Veltz, 1996) et se trouve réutilisé au service de politiques publiques qui en appellent à une nouvelle forme de « gouvernance territoriale ». Bien qu'elle aborde le thème de l'agglomération et de la concentration spatiale des activités, notre approche s'éloigne des préoccupations des sciences régionales et urbaines qui traitent de « systèmes productifs locaux » et de « systèmes locaux d'innovation », autrement dit de l'articulation entre collectivités territoriales, entreprises, recherche et enseignement supérieur dans le milieu local. En effet, pour appréhender la structuration et l'organisation du système scientifique au niveau mondial, nous évacuerons progressivement la question des interactions intra-locales pour nous concentrer sur l'articulation entre le niveau local que nous aurons à cœur de définir à travers la conception des « agglomérations scientifiques », le niveau national, les niveaux régionaux intermédiaires et le niveau mondial.

Ce choix méthodologique nous conduit à formuler une troisième et dernière proposition. Elle découle du fait que nous avons choisi de délaisser la question de l'écologie des lieux pour lui préférer celle de leur situation, c'est-à-dire, de leur localisation relative à un moment donné (Pumain & Saint-Julien, 1997, p. 32). Traditionnellement, la situation des lieux peut être saisie à l'aide de coordonnées géographiques qui sont reportées sur une carte. Nous montrerons que, dans certains cas, une autre manière d'analyser la position des lieux scientifiques est pertinente. Il s'agit, au lieu de s'en tenir à la distance physique qui les sépare, de mesurer l'intensité de leurs relations à partir du volume de colla-

borations scientifiques qu'entretiennent les chercheurs qui s'y trouvent. Ce faisant, c'est l'intensité des collaborations qui sert de métrique pour mesurer la distance séparant les lieux de science. La réflexion sur la représentation des résultats obtenus à partir de ce type d'analyses de réseaux demande un soin particulier. La formalisation imposée par cette approche s'appuie sur la théorie des graphes, une branche des mathématiques qui trouve des applications dans de nombreux domaines. En sociologie, elle est au fondement de l'analyse des réseaux sociaux. En géographie, elle a longtemps été utilisée pour l'analyse des réseaux d'infrastructure et de transport, mais à mesure que la géographie poursuit son intégration dans l'univers des sciences sociales, elle tend à s'appliquer également à l'étude géographique de phénomènes sociaux. L'une des manifestations récentes de cette évolution est la création du [groupe fmnr](#) (flux, matrice, réseau) en France en 2010 qui défend l'importation des méthodes d'analyse de graphes non-planaires en géographie. Les travaux et la dynamique de ce groupe ont été d'un apport non négligeable pour la formalisation de cette thèse.

Cette thèse présente enfin la particularité de ne pas être un produit isolé. Elle a débuté dans le contexte d'un programme de recherche de grande ampleur financé par l'Agence Nationale de la Recherche (ANR) de 2010 à 2013 : [le programme « Géosciences »](#). Ce programme, coordonné par Michel Grossetti et Denis Eckert depuis l'Université de Toulouse, a été l'occasion de rassembler historiens, sociologues, anthropologues et géographes autour du thème de la géographie des sciences. Il nous a permis de nous confronter à l'étendue des problèmes qu'une géographie des sciences serait en mesure de traiter. La question de comprendre en quoi un savoir-faire et une pensée géographiques pourraient alimenter et enrichir les études sociales des sciences a trouvé deux réponses lors de cette expérience interdisciplinaire.

D'une part, les géographes sont les chercheurs les plus disposés à prendre des unités spatiales pour unité élémentaire d'analyse (Pumain & Saint-Julien, 1997, p. 13). Pour simplifier, le sociologue et l'anthropologue, voire l'historien, voudront plutôt raisonner en retenant l'individu ou l'humain comme unité élémentaire d'analyse et l'économiste en retenant le ménage, l'entreprise ou l'organisation. Cette disposition n'est pas étrangère au rapport que les géographes entretiennent avec la cartographie. D'ailleurs, la sensibilité des géographes à ce type de représentations graphiques est une spécificité qui trouve toute sa force dans l'interdisciplinarité. Habitué à agréger l'information pour être en mesure de la représenter de façon lisible, le géographe ne recule pas devant l'idée d'envisager les entités spatiales comme des objets ayant des relations, quitte à se confronter au risque de « l'erreur écologique » (*ibid.*, p. 36-37). Cette vision est choquante d'un point de vue sociologique car ce n'est pas l'entité spatiale qui agit et se met en relation mais les individus agissants qu'elle contient. Cette question pourrait certainement connaître de plus

amples développements (dans la mesure où certains pourrait objecter que l'entité spatiale est elle-même agissante) mais il n'est pas dans notre intention d'aller au-delà. Retenons qu'il paraît essentiel pour la démarche géographique de se détacher de la finesse de l'analyse sociologique pour embrasser d'un seul regard l'ensemble de l'information disponible sur un phénomène. C'est ce que l'opération d'agrégation des données au niveau d'entités spatiales permet de faire. De la même manière, les sociologues ne reculent pas devant l'idée d'agréger l'information au niveau de grandes variables telles que le genre ou l'âge. Si procéder de la sorte avec des variables géographiques n'est pas inconcevable ou interdit par la discipline sociologique, cela reste une pratique moins spontanée pour la profession. Évidemment, nul besoin de s'encombrer de traitements géographiques quand la question de recherche ne le justifie pas, mais ces traitements offrent un bon moyen d'échapper à ce que Denise Pumain qualifie d'« erreur atomiste » (par opposition à « l'erreur écologique ») qui consiste à envisager l'individu comme un être isolé, détaché du sol, et dont le comportement serait indépendant de la localisation géographique (*loc. cit.*).

D'autre part, les géographes ont une certaine familiarité avec les jeux d'échelle, comme le prouve leur intérêt pour les fractales. Ils ont l'habitude d'agréger l'information à différents niveaux, problème parfois difficile, qui « invite à observer des règles précises, qui contrôlent la perte d'information lors de la généralisation » parce que comme l'indique Denise Pumain : « les structures spatiales sont fondamentalement hétérogènes. » (Pumain & Saint-Julien, 1997, p. 39). Or, les chercheurs en sciences sociales (mais pas seulement) sont souvent confrontés à des phénomènes concevables à plusieurs niveaux d'analyse. Ils doivent parfois composer avec le fait que les tendances d'un même phénomène peuvent aller dans différentes directions selon le niveau d'observation. Il arrive aussi que les outils et les méthodes, permettant de traiter l'information à un niveau donné, puissent être inenvisageables à un autre niveau. Il se trouve que dans le cas des activités scientifiques, nous avons affaire à un phénomène qui se manifeste sous des modalités différentes suivant les époques, les lieux et les sujets de recherche, et qui est soumis à différentes logiques organisationnelles et finalités. De ce fait, au même titre qu'il est pertinent de mobiliser des spécialistes du temps et de la société pour étudier la dimension sociale des sciences, le regard des spécialistes de l'espace géographique se justifie. En particulier, au niveau individuel, un chercheur peut appartenir à la fois à des organisations scientifiques locales, régionales, nationales, transnationales et internationales dont les intérêts peuvent diverger, sans compter qu'il est confronté au fait que les pratiques et les valeurs peuvent varier dans le temps et l'espace, y compris au sein de sa propre spécialité de recherche. Ces spécificités sont bien connues en *science studies*, mais saisir leur manifestation dans

l'espace géographique pose des problèmes méthodologiques et présente des enjeux théoriques qui sont au cœur de ce travail.

Au bout du compte, cette thèse sera l'occasion de soulever trois grandes questions :

1. Est-ce que le *Web of Science* est une source d'information pertinente pour saisir la manifestation des activités scientifiques dans l'espace géographique mondial, et si oui, quel aspect des activités scientifiques cette base de données permet-elle de capturer et comment ?
2. Quelles sont les grandes tendances récentes qui affectent les activités scientifiques dans le monde et peut-on parler de mondialisation des activités scientifiques ?
3. Que peut apporter une entrée géographique pour étudier les pratiques scientifiques elles-mêmes et, plus précisément, la constitution et la dynamique d'un collectif scientifique dans le temps et l'espace, en l'occurrence la communauté de la réparation de l'ADN ?

Avant d'en venir à la formulation précise de ces questions, aux moyens que nous avons retenus pour y répondre et à la restitution des résultats, il est apparu nécessaire de définir un socle pour une géographie des sciences.

La première partie est donc l'occasion de réfléchir aux évolutions de la géographie, aux évolutions des études sociales des sciences et des sciences en général, et à la pertinence d'une approche géographique des sciences compte tenu des tendances respectives qui traversent la géographie, les sciences sociales, les études sociales des sciences et les sciences. À l'issue de cette revue de la littérature, nous mettrons à l'épreuve l'idée que les publications visibles dans la source de référence, le *Web of Science*, sont le reflet d'un « monde scientifique », celui de la science « visible ». Nous montrerons que ce monde n'est pas et ne peut pas être représentatif de l'intégralité des travaux de recherche effectués sur le globe terrestre. Pour autant, cet ensemble est un objet intéressant à plusieurs titres : parce que la volonté d'y participer contribue à standardiser l'activité de publication au niveau mondial ; parce qu'alors qu'en même temps qu'il facilite la circulation des savoirs entre disciplines et entre pays, il peut être utilisé comme instrument de hiérarchisation ; parce que même s'il peut servir à évaluer l'activité des scientifiques, il est aussi une plateforme d'observation depuis laquelle on peut repérer l'impact des coupes budgétaires étatiques ou des crises sociétales sur le développement scientifique, mais également les dynamiques propres à l'activité scientifique elle-même.

La deuxième partie est fondée sur la volonté d'analyser le « monde scientifique » reflété par le *Web of Science* et ses dynamiques globales. Pour ce faire, nous présentons la méthode de bibliométrie spatiale que nous avons mise au point, en équipe, dans le cadre

du programme de recherche « Géoscience ». Après avoir justifié nos choix méthodologiques en matière de niveau de résolution, de traitement des données, de visualisation et d'analyse, et avoir discuté de leurs conséquences, nous en venons à la détection des grandes tendances contemporaines dans la répartition des activités scientifiques et leur organisation. Plus particulièrement, nous nous intéressons à la manifestation de deux phénomènes : la déconcentration spatiale de l'activité scientifique et la densification des collaborations scientifiques au cours de la dernière décennie (1999-2008).

La dernière partie porte sur la formation d'un collectif scientifique à l'échelle mondiale qui se retrouve autour du thème de la réparation de l'ADN. C'est à la fois l'occasion de remonter dans le temps et de se rapprocher des contenus scientifiques. En étudiant l'émergence de ce collectif au cours des années 1960-1970, nous montrons qu'il existe des logiques d'organisation, un rapport au temps et à l'espace caractéristiques de ce collectif et des questions de recherche qui le mobilisent. En se penchant sur l'une des questions relatives à la réparation de l'ADN qui a suscité l'intérêt dans les années 1990, nous constatons que l'organisation spatiale des scientifiques en action est susceptible d'évoluer d'un stade de recherche à un autre. Il ressortira, à l'issue de cette partie, que la géographie des collectifs de recherche n'est pas figée ou surplombante mais qu'elle est le produit de la participation des individus et de leurs trajectoires. Une étude rapprochée d'un membre du collectif témoignera du fait que les collectifs scientifiques admettent autant de définitions et de délimitation qu'ils comprennent de membres. La représentation que les scientifiques se font de la « communauté scientifique » à laquelle ils appartiennent dépend à la fois de leurs trajectoires géographiques et scientifiques.

PARTIE 1. PENSER LA SCIENCE À L'ÉCHELLE DU MONDE EN GÉOGRAPHIE

CHAPITRE 1. SCIENCES DE LA SCIENCE, SCIENCES SOCIALES ET GÉOGRAPHIE

Dans *Un tout petit monde*, le romancier David Lodge donne une illustration du système scientifique contemporain : « Chaque discipline, et chaque colloque qui lui est consacré, est un monde en soi, mais les disciplines se rassemblent en galaxies, de sorte qu'un voyageur habitué à se déplacer dans l'espace intellectuel (comme, disons, Morris Zapp) peut passer sans difficulté de l'une à l'autre, figurer comme sémiologue à Amsterdam, comme joycien à Zurich et comme narratologue à Vienne. » (Lodge, 1984, p. 341). Le terme de « voyageur » retenu par Lodge mérite toute notre attention : Zapp passe d'une spécialité à une autre aussi aisément qu'il se déplace de villes en villes. Doit-on comprendre que « le voyageur » change de spécialité rien qu'en changeant de ville ? Dans le récit, le chercheur américain Morris Zapp, de tournée estivale en Europe, présente la même communication à chacun des différents colloques où il se rend mais c'est le cadre dans lequel elle se fait qui diffère et potentiellement, sa réception. Dans les faits, quelle que soit la ville dans laquelle il communique, l'Américain choque les Européens avec sa posture « post-structuraliste » selon laquelle « tout décodage est un nouvel encodage » et déclenche des réactions d'incompréhension. Dans ce récit, Lodge suggère subtilement qu'il existe une intrication entre l'espace intellectuel et l'espace géographique. Les spécialistes et les spécialités ne se répartissent pas uniformément dans l'espace géographique, ce qui n'est pas étranger au fait que chaque lieu de production scientifique possède ses caractéristiques, son histoire, ses équipements, son organisation, sa situation, sa culture et sa réputation vis-à-vis des autres lieux scientifiques. C'est aussi ce que l'auteur évoquait déjà avec humour dans le premier volet de la trilogie intitulé *Changement de décor* (Lodge, 1975). Quand un universitaire britannique et un universitaire californien en viennent à occuper pour quelques mois le poste de l'autre dans le cadre d'un échange organisé par leurs institutions respectives, le dépaysement est complet tant sur le plan personnel que professionnel. Avec ses fictions sur le milieu universitaire, Lodge met le doigt sur des questions qui sont familières aux spécialistes des *science studies*.

1. La science comme fait social

« *Sans institution, la science n'est qu'une forme de passe-temps privé.* » (Gingras, 2013, p. 30)

1.1. Social versus cognitif ?

Sous l'étiquette *science studies* se retrouve en 2015 une grande variété de questions de recherche et d'approches disciplinaires ; ceci dit, l'expansion de ce domaine est venue pour une grande part de la sociologie dans la deuxième moitié du XX^e siècle. La sociologie des sciences et des connaissances scientifiques, dont l'un des moments fondateurs fut la thèse de Robert K. Merton (parue en 1938 dans la revue d'histoire des sciences *Osiris*), est pourtant restée assez confidentielle jusqu'aux années 1980, voire 1990 (Merton, 1938 ; Dubois, 2001). Son tour de force a été de rendre acceptable l'idée que le fait scientifique, au même titre que le fait politique, religieux et économique, est un fait social. Pouvoir aborder le fait scientifique comme un fait social est indispensable pour nourrir une réflexion sur la science en sciences sociales ; encore fallait-il s'accorder sur ce qu'il y a de social dans le concept de science. Autrement dit, est-ce que le caractère social de la science s'arrête à la dimension pratique et institutionnelle de l'activité scientifique ? Il reste d'usage en sociologie de distinguer ce qui tient à l'institution scientifique, ses normes et fonctionnements, et ce qui relève du cognitif, les idées et leur inscription dans un environnement donné et situé. Des méthodes, concepts et théories ont été forgés par plusieurs générations de spécialistes des sciences pour tenir compte de ces deux aspects et, au bout du compte, de leur indissociabilité (Zuckerman, 1988). Les évolutions dans ce domaine ne se sont pas toujours faites en douceur et l'apport des différentes disciplines qui y ont contribué (histoire, philosophie, anthropologie) est complexe à analyser. Dans ce processus, un ouvrage se démarque et mérite d'être signalé : *La Structure des révolutions scientifiques* du physicien, historien et philosophe Thomas Kuhn (Kuhn, 1962). Son retentissement dans le monde scientifique sort de l'ordinaire, notamment en France où il est traduit pour la première fois en 1972. Il reste l'un des livres les plus cités du XX^e siècle. À titre indicatif, pas moins de 80 000 citations de ce livre étaient recensées dans *Google Scholar* en juillet 2015. Cet ouvrage présente une vision du progrès scientifique qui tranche avec la représentation traditionnelle : plutôt que de suivre un développement linéaire et cumulatif, le savoir scientifique s'inscrit dans des cadres de pensée ou « paradigmes » qui peuvent entrer en contradiction, voire en conflit. L'idée marquante véhiculée par ce travail, largement discutée depuis, est l'idée que l'on peut détecter des facteurs sociaux qui permettent à un

paradigme de l'emporter sur un autre en situation de controverse. Pour formuler cette théorie, Kuhn reconnaît dans sa préface l'influence qu'a eue sur lui l'ouvrage du sociologue allemand Ludwick Fleck : *Genèse et développement d'un fait scientifique*, passé relativement inaperçu lors de sa parution en 1935 (Fleck, 1935 cité par Kuhn, 1962). Pour Fleck, le fait cognitif est un fait social et la pensée scientifique est sensible au « style de pensée » qui imprègne la société. En cela, la théorie de Fleck est encore plus radicale que celle de Kuhn qui n'emprunte à Fleck que sa notion de « collectif de pensée » (Braunstein, 2003). Grâce à Kuhn, l'ouvrage de Fleck a bénéficié d'une seconde vie et beaucoup de sociologues des sciences s'en sont emparés. Dans le langage des scientomètres, le livre de Fleck est une « belle au bois dormant » qui a été réveillée par son « prince », *La Structure des révolutions scientifiques* (à propos des « belles au bois dormant », voir par exemple Braun, Glänzel, & Schubert, 2010).

Thomas Kuhn n'a pas été satisfait de la réception de son livre parmi les sociologues, notamment parce que lui-même n'était pas allé jusqu'à défendre l'idée qu'un « style de pensée » puisse influencer les contenus scientifiques, considérant que ce sont uniquement les changements de paradigme qui peuvent en être affectés. Néanmoins, *La Structure des révolutions scientifiques* a servi de point de départ au développement de la sociologie des connaissances scientifiques (ou *Sociology of Scientific Knowledge*). Cette sociologie rompt avec la sociologie des institutions scientifiques, dite « externaliste » ou « fonctionnaliste », pour s'intéresser aux déterminants sociaux des contenus scientifiques. L'approche « internaliste » la plus extrême, et dont Kuhn se défend, s'appuie sur un ensemble de principes formulés dans le cadre d'un programme de recherche constructiviste baptisé par ses promoteurs le « programme fort » (Bloor, 1976), par opposition à un programme de recherche jugé « faible », celui des philosophes des sciences Thomas Kuhn et Imre Lakatos. Parmi les principes du programme fort, le principe de symétrie est le plus controversé qui incite le sociologue à traiter sur un pied d'égalité les succès et les échecs en science sans se préoccuper de valeurs comme la « vérité ». Pour Thomas Kuhn comme pour le français Raymond Boudon, le programme fort implique une posture relativiste qui présente un danger pour l'intégrité du savoir scientifique (Boudon, 2008 ; Kuhn, 2000). Une telle posture questionne « l'autonomie » de la science, et en particulier « l'objectivité » et « l'universalité » des énoncés scientifiques puisqu'elle met en évidence les particularismes et les contingences dont ils dépendent. Or, pour Kuhn, et plus généralement pour les rationalistes, il y a des logiques internes à la science qui en garantissent l'efficacité en tout temps et en tous lieux. Cela dit, rien n'empêche de penser que ces logiques internes qui relèvent du fait cognitif soient également sociales. En effet, Fleck écrit : « Celui qui (...) considère le facteur social comme un *malum necessarium*, comme une insuffisance malheureusement constitutive du genre humain et qu'il faut se faire un

devoir de combattre, celui-ci méconnaît qu'il n'y a vraiment aucun acte cognitif qui ne soit possible sans facteur social, que l'expression "acte cognitif" n'a de signification que lorsqu'elle est en relation avec un collectif de pensée. » (Fleck, 1935 cité par Granjou & Peerbaye, 2011). Si l'on admet que les logiques internes à la science sont également sociales, le principe de symétrie n'échoue à en rendre compte que si l'on se limite à expliquer le succès ou l'échec des énoncés scientifiques à la lumière de facteurs sociaux extérieurs à la science, sans considération, par exemple, pour la rationalité des énoncés.

Dès lors, ce n'est pas le principe de symétrie qui présente un danger pour la science mais le fait d'y associer un scepticisme vis-à-vis de l'autonomie et du bien-fondé des énoncés promus au rang de vérités scientifiques. Or, dans les cas où elle se traduit par un refus de hiérarchiser les idées, les savoirs et les valeurs, la posture relativiste autorise le nihilisme mais elle ne l'impose pas. Les controverses apparues dans les années 1990 autour des *cultural studies* ont été typiques des dérives de cette approche et c'est parce qu'en questionnant le concept d'« objectivité », la sociologie des connaissances scientifiques s'est permise de remettre en cause la « vérité absolue », qu'elle a été tenue responsable de ces dérives. Alors qu'en 1996, le physicien Alan Sokal tâche de discréditer cette sociologie, il contribue involontairement à la populariser (Bourdieu, 2002b, p. 10 ; Dubois, 2001, Chapitre 6 et 7). Bien que suite à cette polémique, beaucoup de chercheurs se soient défendus d'être « relativistes », le fait d'avoir caricaturé cette sociologie ne lui a pas permis d'en disqualifier l'apport. Ne se réduisant pas à une posture « relativiste », la sociologie des sciences n'a pas faibli devant les attaques. Elle a su résister car, dans une large mesure, son intention n'était pas de remettre en cause le bien-fondé de la démarche scientifique : le fait que des facteurs sociaux extérieurs à la science influencent la pratique et les contenus scientifiques n'empêche pas qu'il existe une raison propre à la science et des logiques sociales internes qui garantissent son efficacité. Si la science – à la fois les institutions et les connaissances scientifiques – est un fait social, elle ne saurait être entièrement déterminée par des facteurs sociaux qui lui sont extérieurs. Cela explique que le thème de l'« autonomie » de la science soit si important, notamment chez Pierre Bourdieu pour qui l'autonomie du « champ scientifique » est au fondement de son intégrité (*ibid.*, p. 91-109).

1.2. Changements de « programmes »

Ainsi, la pensée de Kuhn a largement été discutée, interprétée et approfondie jusqu'à aujourd'hui, et beaucoup diraient « pour le meilleur et pour le pire ». En particulier et indépendamment de la sociologie, un ensemble de travaux épistémologiques a cherché à saisir des moments d'évolution scientifique à l'aide d'une grille de lecture kuhnienne. Bien qu'heuristique, la mobilisation de cette grille de lecture a depuis été jugée

trop systématique et simpliste. Selon les spécialistes de l'œuvre de Kuhn, il faudrait avoir à l'esprit que cette grille de lecture n'est pas valable et valide pour cerner indifféremment toutes les évolutions scientifiques ; en effet, le concept de « révolutions scientifiques » a été mis au point sur le modèle de la physique. La question de la pertinence de la grille kuhnienne s'est toutefois posée pour la géographie. Cette dernière, plus encore en France que dans le monde anglo-américain, est restée à l'écart des études sur les sciences. Pourtant, elle a connu d'importants bouleversements épistémologiques autour des années 1970 qui furent, comme la « révolution quantitative », alimentés par des discours s'inspirant du modèle kuhnien (Berry, 1973 ; Haggett & Chorley, 1967). Après coup, des géographes ont réalisé que le modèle kuhnien n'était pas adapté pour rendre compte de l'évolution de leur discipline (Wheeler, 1982 ; Mair, 1986). Dans l'espace anglophone, Ron Johnston montre qu'au lieu de voir un paradigme dominant l'emporter, par exemple celui de l'analyse spatiale, la géographie est devenue plurielle (Johnston, 2006). Olivier Orain tire des conclusions similaires pour le cas français (Orain, 2006). Néanmoins, il considère que la notion de paradigme reste adaptée pour décrire l'âge classique de la géographie qui s'étend de la Belle Époque jusqu'aux années 1950. Cette période est dominée par l'« école vidalienne » et laisse peu de place, en France, à d'autres conceptions et manières de faire de la géographie humaine.

Finalement, il est intéressant de noter que d'une part l'œuvre de Kuhn a eu des répercussions immédiates sur les contenus scientifiques puisqu'elle a été mobilisée ou même instrumentalisée, notamment par les géographes, pour faire évoluer leur propre discipline et d'autre part, qu'elle sert régulièrement de matrice ou de grille de lecture pour les spécialistes des sciences qui tâchent d'analyser des phénomènes de développements scientifiques. À cette seconde fin, le philosophe Adi Ophir et l'historien Steven Shapin considèrent que « le travail de Kuhn a offert des ressources pour un programme de recherche capable de relier l'ontologie et l'épistémologie aux spécificités des différentes positions sur la carte sociale. »¹ (Ophir & Shapin, 1991). De l'avis d'Ophir et Shapin, *la Structure des révolutions scientifiques* participe de l'adoption d'un nouveau programme de recherche ou paradigme pour saisir les logiques d'évolution scientifique en *science studies*.

La pertinence de ce programme – qu'une géographie des sciences, nous y reviendrons, est en mesure d'intégrer à son agenda – ne tient pas, loin de là, au contenu précis du livre de Kuhn, mais à son inscription dans un contexte favorable à l'appréhension de la vérité scientifique, voire de la réalité qu'il suppose. Comme nous l'avons déjà évoqué, les idées véhiculées par cet ouvrage ne sont pas inédites et Kuhn n'est ni leur seul porte-

¹ L'original : « Thus, Kuhn's work offered resources for a program of research that related ontology and epistemology to the particularities of different places on the social map. Inspect different social focus and you will discover different patterns of scientific judgment. » (Ophir & Shapin, 1991).

parole, ni le plus radical. Cependant, leur formulation dans *la Structure des révolutions scientifiques* arrive au moment de la remise en cause du positivisme logique, qui avait déjà été amorcée entre autres par les philosophes austro-hongrois Karl Popper et Michael Polanyi (Polanyi cité par Thomas S. Kuhn, 1990 ; Popper cité par Ophir & Shapin, 1991). Dans le contexte de la guerre froide, une vision constructiviste de la connaissance se développe dans le monde dit « libre » : les théories ne sont pas vraies en soi mais peuvent être déclarées vraies dans un cadre de pensée donné. Or, le postulat de la contingence du progrès scientifique, s'il est bien compatible avec une vision constructiviste de la connaissance, ne l'est pas avec une vision positiviste. Dès lors, plutôt que d'être responsable d'une « rupture paradigmatique » dans le milieu scientifique, *la Structure des révolutions scientifiques* s'inscrit dans un processus de défiance à l'égard du « progrès » permettant de repenser, comme le suggèrent Ophir et Shapin, le lien entre l'ontologie et l'épistémologie en tenant compte de la subjectivité des chercheurs et du cadre de pensée qui est le leur (*ibid.*). En sociologie des sciences, on l'a déjà évoqué, l'approche constructiviste se caractérise par la prise en compte des contenus scientifiques dans l'analyse, puisqu'à partir du moment où ces contenus sont aussi des construits sociaux, une sociologie de la construction des contenus scientifiques est possible.

Ce changement d'attitude par rapport à la connaissance touche toutes les sciences sociales. Celles-ci sont en proie à une forte expansion après les mouvements étudiants de la fin des années 1960 sur fond de massification scolaire et de déconcentration spatiale de l'enseignement supérieur et de la recherche. Progressivement les effectifs de chercheurs augmentent ce qui laisse la place au développement d'une grande variété de recherches et de courants épistémologiques. À partir des années 1990, on ne compte plus les branches ou sous-spécialités en sociologie, histoire, géographie, anthropologie ... Cette situation est favorable à l'idée qu'il n'y a pas un seul cadre théorique qui soit supérieur aux autres pour appréhender le réel et progresser. Par conséquent – et cette idée d'un espace scientifique pacifié est plutôt post-kuhnienne – il ne sert à rien de les opposer puisqu'à eux tous, ils offrent différentes représentations et réalités. Pour certains observateurs, ces évolutions intellectuelles sont caractéristiques du passage de la « modernité » à la « postmodernité », du « positivisme » au « post-positivisme » ou encore du « structuralisme » au « post-structuralisme ». Comme le suggère Bourdieu (2002b, *op. cit.*), des facteurs conjoncturels peuvent expliquer ces transitions et particulièrement l'accession d'une nouvelle génération de chercheurs souhaitant s'inscrire en rupture vis-à-vis de leurs prédécesseurs. Sans entrer dans le détail de ces postures théoriques, il semble que l'appréhension du réel et de l'entendement qui tend alors à s'imposer dérive de la dialectique transcendantale d'Emmanuel Kant (c'est entre autres l'avis de Richard Rorty, 1979 cité par Ian Hacking, 1980). Si l'on distingue le réel « en soi » (noumène) et le réel perçu (phénomène), et si

L'on admet que le premier est inatteignable, y compris par l'entendement, alors il faut aussi admettre que les structures sociales ou les lois d'organisation humaines que l'on peut abstraire de la réalité en s'appuyant sur un raisonnement scientifique n'existent pas « en soi » (Bitbol, 1998). Ceci dit, ces structures ou ces lois sont des clefs d'interprétation, des images qui influencent ou informent notre perception du réel.

Alors que l'ouvrage de Thomas Kuhn est souvent mobilisé pour rendre compte et expliquer les mutations qui caractérisent la vie intellectuelle de la dernière partie du XX^e siècle d'une part et les *science studies* d'autre part, les lignes qui précèdent incitent à nuancer son empreinte. Victime – non consentante – de son succès, ce livre n'est pas un manifeste pour l'approche constructiviste des sciences et de la connaissance bien qu'il en soit devenu emblématique. Le fait est qu'il est plus commode d'attribuer à cet ouvrage un rôle décisif que de prétendre saisir, dans leur intégralité, les phénomènes qui ont touché au fait scientifique et intellectuel à partir des années 1960. Sans qu'il soit ici question d'opter pour l'une ou l'autre des possibilités, il convient finalement d'admettre que cet ouvrage n'est pas une référence suffisante, mais qu'il a participé à la remise en cause, dans les années 1970, des cadres de pensée de la génération active dans l'immédiat après-guerre. La vision de la science et de la connaissance s'en est trouvée modifiée si bien que la science – comme objet social – à la fois pratiques et contenus – s'est elle-même transformée. Si l'on convient que la pratique scientifique (le tangible) est indissociable des contenus scientifiques, alors il est vain de se demander ce qui est premier : l'évolution du concept de science (le cognitif) ou bien les mutations de la science en pratique (le social). Ce qui est certain, c'est que la vision dominante de la science a évolué, que la science a évolué et qu'évidemment dans le même temps, la science de la science ou, plus précisément, le programme de recherche des *science studies* a évolué. L'idée cruciale qui sous-tend ces mutations est que l'intrication entre le savoir et la pratique, entre l'entendement et le sensible est indémêlable. Cela implique que le réel en lui-même est inatteignable et au-delà, que le fait cognitif (les idées, la connaissance, le savoir scientifique) est tout aussi social que le fait pratique.

C'est en mobilisant cette approche que des géographes en sont venus à accepter de considérer la géographie comme une science sociale (Di Méo, 2008), et cela y compris par-delà de la détermination de plusieurs géographes, à la fin des années 1980, à participer au développement d'une sous-catégorie de la discipline, la géographie sociale (Frémont, 1986 ; détermination critiquée par R. Brunet, 1986). La géographie est une science ou discipline qui comprend l'ambition de saisir le monde terrestre compte tenu de son rapport à la société. À partir du moment où l'on adhère à l'idée qu'il n'est pas possible d'appréhender le monde indépendamment de la société parce que la perception et par conséquent la représentation que l'on peut s'en faire est nécessairement subjective et in-

complète ; et ensuite, que cette perception et cette représentation sont sociales ; alors il devient pertinent de considérer la géographie comme une science sociale. Dans le même temps, puisque la science peut être appréhendée comme un objet social, comme l'un des grands faits sociaux des sociétés humaines, elle est un objet potentiel pour la géographie humaine (Livingstone, 2003). Dans les sections qui suivent, nous revenons sur la place de la géographie dans l'espace des sciences sociales et questionnons l'intégration possible de la « science » comme objet d'étude. Plus particulièrement, nous entendons montrer qu'en raison de son autonomie vis-à-vis des autres faits sociaux, la science mérite d'être abordée en géographie, non pas seulement par le biais de ses relations aux autres aspects de la vie sociale (économiques, politiques ou religieux) mais pour elle-même. Il s'agit en somme de démontrer qu'une branche de recherche tournée vers les rapports entre sciences et espace géographique est légitime et présente un intérêt pour la discipline de même que pour les *science studies*. À l'issue de ce premier chapitre, nous aurons posé que la géographie ayant elle-même été traversée par de profondes mutations épistémologiques dont elle est ressortie « plurielle » (Orain, *op. cit.*), elle apparaît mûre pour accueillir ce projet qui nécessite un rapprochement avec les *science studies*.

2. La géographie comme science humaine et sociale

« Tout fait culturel naît dans un creuset social et se constitue dans le tissu des rapports sociaux et spatiaux ; même si, en retour, les contenus culturels façonnent les contextes sociaux et spatiaux qui les produisent et les expriment. » (Di Méo, 2008)

2.1. Le « tournant culturel » en géographie

Pour que des chercheurs en géographie s'aventurent dans l'univers des *science studies*, il fallait à la fois qu'ils soient prêts à considérer la science comme un fait social et culturel et leur propre discipline comme une discipline pouvant prendre comme objet de tels faits, y compris ceux qui touchent au domaine des idées et n'impactent pas directement l'environnement physique ou matériel. Justement, à la fin du XX^e siècle, le mouvement intellectuel qualifié de « tournant culturel » a renouvelé le regard d'une partie des géographes, en particulier ceux qui se retrouvent dans la géographie économique, sur leur activité. Selon Trevor Barnes, l'ouvrage de Derek Gregory, *Ideology, Science and Human Geography*, paru en 1979, a joué un rôle important dans ce processus. Il aurait contribué à pousser la géographie anglo-américaine dans l'espace des sciences sociales (Gregory, 1979

cité par Barnes, 2001). Dans son article portant sur ce tournant en géographie économique, Barnes amène le sujet par son expérience personnelle de Britannique parti étudier de l'autre côté de l'Atlantique. Il reconnaît ainsi la part de subjectivité qu'il y a dans son appréhension de la discipline géographique. En particulier, il est conscient que le livre de Gregory fut avant tout un événement pour la géographie anglo-américaine. Le fait que Barnes fasse preuve de réflexivité dans cet article est cohérent avec le propos qu'il défend. Selon lui, les travaux issus de la « révolution quantitative » en géographie étaient, jusqu'au « tournant culturel », alimentés par une argumentation, une méthodologie et une posture qu'il qualifie d'« épistémologique ». Or, Barnes considère que le livre de Gregory a participé d'un tournant favorable au développement d'un rapport à la théorie qu'il qualifie d'« herméneutique » et qui est propice à la réflexivité. Pour lui, le mode de théorisation est « épistémologique » quand son ambition est de donner à voir le réel objectivement, le seul moyen d'y parvenir étant de s'appuyer sur un langage et des méthodes empruntés aux sciences dites « exactes » comme la logique, les mathématiques ou les sciences physiques et d'ignorer l'impact de la position spatio-temporelle du chercheur sur ses réflexions et son travail. En revanche, la théorisation est « herméneutique » quand son objet, quels que soient le vocabulaire ou les méthodes utilisés, porte sur les interprétations du réel en acceptant la part de subjectivité propre à toute théorie. Cette disposition d'esprit vis-à-vis de la théorie justifie l'adoption d'une posture réflexive : « théoriser est une activité sociale comme une autre. Les théories ne tombent pas du ciel ; elles reflètent le contexte social local et les relations inégales de pouvoir qui s'y manifestent. Dès lors, théoriser requiert une mise en situation autrement dit, il convient de mettre en avant ses propres intérêts, sa localisation et son identité ainsi que celles des autres » (*ibid.*). À notre sens, il ne s'agit pas de condamner la première posture mais de montrer que d'autres postures sont tenables. En fait, s'il est louable de prétendre à l'objectivité, il faut être conscient que l'objectivisation totale est une vue de l'esprit (Avramides, 2006).

En France, la recension de l'ouvrage de Gregory a été faite par Antoine Bailly dans l'*Espace Géographique* en 1979 (Bailly, 1979). Avec Paul Claval, Bailly a importé dans le monde de la géographie francophone, des fragments de géographie anglo-américaine avec pour projet de poser les fondements d'une « Nouvelle Géographie » (Claval, 1977). Ouverte aux méthodes quantitatives, cette géographie ouvrait aussi la porte aux représentations et aux effets culturels tout en restant sensible à la critique marxiste. En s'appuyant sur Gregory, Bailly affirme : « la géographie ne peut rester une science spatiale, une discipline de la seule logique spatiale. Elle doit s'ouvrir à d'autres voies, structuralistes, phénoménologiques, critiques. En ce sens, l'intégration des systèmes humains et physiques ne constitue pas un problème épistémologique : les deux sont liés par la pratique sociale. La géographie n'existe qu'en tant que sous-ensemble de l'herméneutique, c'est-à-dire de

l'étude de l'interprétation des relations générales de l'homme au monde. » (Bailly, *op. cit.*). De l'avis d'Olivier Orain, en proposant une géographie des représentations avec Armand Frémont, Bailly a participé à promouvoir une posture géographique plus constructiviste (Orain, 2014). Toutefois, plutôt que d'« un basculement de la communauté des géographes du réalisme au constructivisme », Orain parle de co-existence (*ibid.*). Avec l'avènement d'un « constructivisme géographique » dont on peut discuter l'origine (il serait réducteur de l'attribuer uniquement à l'influence des géographes américains post-modernes), il n'y a pas qu'une seule voie pour accéder à la scientificité mais plusieurs, parce qu'il existe autant de réalités que de cadres pour saisir le réel. Dans cette mouvance, la géographie contemporaine a pu se diversifier et devenir « plurielle » (Orain, 2006), sans perdre, pour autant, son statut de discipline scientifique. À l'heure de la « géographie des territoires » (Orain, 2014, *op. cit.*) il n'est pas nécessaire d'employer des méthodes issues des sciences « exactes » et de chercher à énoncer des lois pour faire acte de géographe. Les approches plus qualitatives, comme les méthodes d'entretien, offrent des moyens tout à fait appropriés pour faire de la « nouvelle » géographie (Claval, *op. cit.*). Il convient désormais d'admettre que toute démarche qui décrit, analyse, représente et rend compte des interprétations du réel en un temps et un lieu donné est acceptable en géographie. C'est ainsi, par exemple, que Gérard Dussouy se faisait récemment le défenseur d'une approche pragmatique, inspirée par le cercle herméneutique, en géopolitique (Dussouy, 2011).

Sans qu'il faille exagérer l'impact de l'ouvrage de Gregory sur l'ensemble de la discipline géographique (*Google Scholar* en recense 564 citations en juillet 2015), *Ideology, Science and Human Geography* a sans doute contribué à conforter la place de la géographie dans l'espace des sciences humaines et sociales. En ce sens, il est porteur d'un projet et son ton est volontiers performatif. L'intention est d'inscrire la géographie dans un vaste univers d'études des faits sociaux et des sociétés humaines qui est celui des sciences humaines et sociales. Comme science spécialisée dans le traitement de la dimension spatiale des phénomènes, la géographie a une carte à jouer. Or, dans les années 1970, la réussite de ce projet n'avait rien d'une évidence, notamment en France, où la géographie a longtemps été reléguée au rang de discipline scolaire (Rhein, 1982). Suscitant le désarroi des géographes, Jean Piaget ne prend même pas la peine de l'évoquer en 1967 dans *Logique et connaissance scientifique* (Orain, 2006 ; Raffestin & Lévy, 1998). Aujourd'hui encore, la géographie n'est pas la première des disciplines qui vient à l'esprit lorsqu'il est question de sciences humaines et sociales (Wallerstein, 1982 n'en parle pas non plus dans sa conférence sur le concept de développement en sciences sociales). Une partie des géographes, en particulier ceux qui se réclament de la géographie physique, restent plus proches des sciences naturelles et du vivant que des sciences humaines et sociales. À l'opposé, les géographes français favorables à l'idée que la géographie traite de faits sociaux, et en particu-

lier de leur déclinaison dans l'espace géographique, sont ceux que Claval considère comme participants à la « Nouvelle Géographie » (Claval, 1972, 1977). Cette matrice disciplinaire a en commun avec la « Nouvelle Histoire » française, d'accueillir une grande variété d'objets d'étude et d'approches scientifiques. Pourtant, toutes deux sont restées relativement indifférentes au fait scientifique. Persuadée que la géographie pourrait aborder la science et les activités scientifiques de façon complète – c'est-à-dire en absorbant les questionnements des *science studies* – nous proposons des pistes, dans la suite de cet argumentaire, qui pourraient expliquer la rareté des travaux de géographie ayant, jusqu'ici, traités de la science et des activités scientifiques.

Dans un article paru en 2003 dans *Cahiers d'Histoire* ayant un objectif analogue, Jérôme Lamy montre que « la science » n'a pas été un objet tellement abordé par les historiens français de la Nouvelle Histoire (Lamy, 2003). Il remarque qu'en France, l'histoire des sciences – qui, contrairement à la géographie des sciences est une branche de recherche déjà bien développée – est surtout faite par des non-historiens. Les historiens de formation ont volontiers laissé cet objet aux philosophes, semble-t-il par crainte de ne pas être compétents ou légitimes pour traiter des monuments de la raison humaine, des grands savants et de leurs idées. Ainsi Lamy intitule-t-il son article : « La science, le continement ignoré des historiens français ? ». À sa lecture, on remarque que pour une partie d'entre eux en tout cas, « les obstacles et les réticences qui ont empêché "la science du passé" de s'intéresser au passé de la science » pourraient tout aussi bien expliquer que la « science de l'espace » se soit encore si peu soucieuse de l'espace de la science. Pour intégrer la science dans leurs problématiques, les historiens ont dû « dépasser à la fois leurs appréhensions nées d'une représentation trop réduite de la science et les clivages disciplinaires qui pèsent sur cet objet de recherche » (Lamy, *ibid.*). D'après nous, cela s'applique également aux géographes qui se sont davantage souciés d'économie que de science pour aborder la société jusqu'aux années 1990. Il faut dire que jusqu'alors la géographie était influencée par l'histoire et par l'économie plutôt que par la philosophie, l'anthropologie ou la sociologie. Cette dernière, un temps critique à l'égard de la géographie, s'en est peu à peu désintéressée pour aller concurrencer la philosophie au tout début du XX^e siècle. Et pour des raisons tant scientifiques qu'institutionnelles, la sociologie a cherché à se positionner par rapport à la philosophie en se saisissant d'objets qui lui étaient propres (Rhein, 1982). Or, « la science », au moins pendant la première partie du XX^e siècle, était un objet réservé à la philosophie (Lamy, *op. cit.*).

Parallèlement, à plusieurs étapes de l'histoire de la discipline géographique en France, la prédilection pour une lecture économique de la société est restée prégnante, même si celle-ci n'a jamais été exclusive, comme le suggère Marie-Claire Robic lorsqu'elle étudie la trajectoire de l'objet « ville » dans la géographie française du début du

XX^e siècle aux années 1950 (Robic, 2003). En particulier, à côté de la tendance sur laquelle nous voulons insister et que Robic qualifie d'« ékonomisme », il existe un fort penchant pour le « naturalisme » dans la géographie française. Mais arrêtons-nous un instant sur son ékonomisme puisqu'alors qu'il n'a jamais manqué de géographes pour étudier les échanges de marchandises, les filières de production et le transport des biens marchands, le tout à différentes échelles (Arrault, 2007), il n'est pas possible d'en dire autant des échanges de connaissances, de l'émergence des domaines de recherche ou encore des déplacements des savants et de leurs spécimens dans l'espace géographique. À notre sens, les activités économiques contrairement aux activités scientifiques ont été beaucoup plus volontiers étudiées par les géographes en raison de la complicité entre la géographie et l'économie. Dès la fin du XIX^e siècle, les tenants de « la doctrine libérale » ont favorisé l'établissement de relations étroites entre la géographie scolaire et l'économie politique (Rhein, 1982 ; Clerc, 2007). Selon Catherine Rhein, cette doctrine « dérivée de l'industrialisme saint-simonien (...) mettait l'accent sur l'importance du développement économique des nations, donc du développement industriel, comme facteur et conséquence du développement moral des nations » (Rhein, 1982). Ainsi, lors de la réforme scolaire de 1872, à laquelle participe Émile Levasseur, la géographie française devient théoriquement liée à l'économie tout en restant dans la dépendance institutionnelle de l'histoire (Clerc, *ibid.*). Le style de pensée qui prévaut est dérivé du « rationalisme scientifique » (Rhein, *op. cit.*). Dans cette pensée, « la civilisation » (au sens du libéraliste Gustave de Molinari, cité par Rhein, *op. cit.*) s'élève grâce aux progrès conjoints de la science et de l'industrie. Le « progrès » permet l'approfondissement de la division du travail, une meilleure exploitation de la nature et donc un accroissement de la richesse, condition du bien-être social. Cette idéologie justifie la poursuite de l'effort colonial français ou « mission civilisatrice » (Petitjean, 2005). Mais, le progrès scientifique semble aller de soi et suivre un ordre logique, vertueux et linéaire qui ne préoccupe que les philosophes (Lamy, *op. cit.*). Le penseur et artiste oublié Gabriel-Jules Delarue (dit Strada) (1821-1902), promoteur d'un scientisme teinté de patriotisme (qui se réclame d'un rationalisme cartésien opposé à Auguste Comte) n'hésite pas à parler d'une « fatalité du progrès » (Strada, 1899, p. 87 ; Landreau, 2012). Dans cette course au progrès, les géographes ont seulement pour mission d'étudier les « forces de la nature » et leur appropriation par les humains (Levasseur cité par Clerc, *ibid.*), mais certainement pas de s'intéresser aux lieux et aux déterminants géographiques dudit « progrès ».

Encore aujourd'hui, quand elle subit l'influence de la doctrine libérale ou même de son pendant critique marxien, la géographie des activités humaines est plus volontiers appréhendée par le biais de leurs logiques économiques, à savoir en vue de la production de richesses (y compris lorsqu'elle pourrait être appréhendée par le biais de leurs logiques

scientifiques, en vue de la production de savoirs). Il faut dire que dans la seconde moitié du siècle, lorsque des géographes s'impliquent dans les chantiers de l'aménagement du territoire, mais aussi dans de grands organismes internationaux participant à la structuration d'un « nouvel ordre mondial », le caractère opérationnel de la discipline tend à s'affirmer et la sensibilité aux doctrines économiques, qui font autorité sur la scène politique, à se confirmer. Les grands thèmes de préoccupation que sont le sous-développement (Lacoste, 1959) et, plus tard, la mondialisation sont avant tout absorbés par la géographie à travers leurs dimension économique (Antheaume *et al.*, 1999). Bien que le concept de « civilisation » soit détrôné par celui de « développement » dans les années 1960 (Wallerstein, 1982), puis de « compétitivité » à partir des années 1980 (Carnaggi & Capello, 2010), la doctrine sous-jacente reste la même, qui établit un lien entre « progrès scientifique », « richesse » et « bien-être » – le dernier dérivant mécaniquement des premiers (Petitjean, 2005). Cette prédilection pour une lecture économique de la société se répercute dans le choix des modèles explicatifs, des méthodes ou encore dans la façon de poser et d'aborder les problèmes.

Ainsi, pour analyser la société, au moins jusqu'aux années 1990, statisticiens, démographes et géographes se réfèrent principalement aux cadres d'analyse développés par des économistes ; en particulier la classification par Jean Fourastié et Colin Clark des activités professionnelles en trois secteurs d'activité : primaire, secondaire, tertiaire. En analyse spatiale, les emprunts faits à l'économie portent également sur la méthode. En effet, tout un pan de l'analyse spatiale cherche à expliquer l'organisation territoriale par des fonctions de minimisation de coûts, à l'aide de modèles et de techniques importés de l'économie (Claval, 1977). Par ailleurs, la lecture économique de la société ou l'effet de « l'encastrement » des sphères sociales et politiques dans la sphère économique (l'expression est de Karl Polanyi) a aussi des répercussions sur la géographie critique. À partir des années 1970, le courant critique en géographie s'appuie sur les travaux de Fernand Braudel et Immanuel Wallerstein et envisage l'espace comme le support de l'inégalité des rapports de production. Sous l'influence de Braudel (Braudel, 1985), le marché est envisagé comme un acteur crucial des faits géohistoriques ; et sous l'influence de Wallerstein (Wallerstein, 2004), le processus de mondialisation est celui de l'expansion du système-monde, d'un réseau d'échanges marchands régi par une logique capitaliste. La géographie de Roger Brunet est, sans conteste, influencée par cette approche marxienne de l'espace (R. Brunet, 2001). En outre, la géohistoire et la géographie systémique ont, parmi leurs ambitions, celle de saisir à plusieurs échelles, en plusieurs endroits et à plusieurs époques, les rouages du système capitaliste et les inégalités dont il se rend responsable (Grataloup, 2010), mais plus généralement les leviers du changement social. Le modèle « centre-périphérie » est souvent mobilisé dans cet esprit. Il faut dire

que la conjoncture économique d'alors se prête bien aux approches critiques puisqu'arrivent les premières crises pétrolières et le chômage de masse. En somme, que ce soit pour les analyser ou les critiquer, les rapports économiques entre les hommes sont centraux en géographie.

Ce rapide tour d'horizon de la discipline pourrait laisser penser que sa dépendance à la doctrine libérale fait figure d'exception. Au contraire, d'après Wallerstein, la doctrine libérale du XIX^e siècle est au fondement des sciences sociales en général. Pour lui, la limite des courants critiques des années 60-70 a justement été de ne pas revenir sur ces prémisses : « Alors que les mouvements de résistance rejetaient tous l'épistémologie de la science sociale dominante, ils n'ont pas contesté son historiographie. Ils acceptèrent sans trop y penser que les deux grands tournants de l'histoire moderne furent la "révolution industrielle" au début du dix-neuvième siècle et la Révolution française. Par le fait même, ils acceptaient la prémisse que la construction d'une économie mondiale capitaliste représentait, en soi, un "progrès", justifiant ainsi implicitement la théorie des "stades de développement" pour chaque société. En abandonnant le terrain historiographique, ils affaiblissaient leur résistance dans le secteur épistémologique. » (Wallerstein, 1982). Ce constat très général devrait sans doute être nuancé, mais c'est à l'appui d'une proposition nouvelle, d'une *via media* (le holisme méthodologique), que Wallerstein s'autorise un tel raccourci.

Aussi, sans prétendre réviser l'historiographie de la discipline et encore moins celle de l'ensemble des sciences sociales, notre insistance sur « l'économisme » de la géographie a pour but d'étayer l'hypothèse centrale du prochain chapitre. D'après nous, l'« économisme » disciplinaire pourrait expliquer que lorsque les activités scientifiques ont été évoquées par des géographes au XX^e siècle, c'était généralement de façon périphérique, à travers leur dépendance aux retombées économiques. Or, en revenant sur l'évolution des sciences de la science, des sciences humaines et sociales et de la géographie, il ressort qu'une géographie des sciences pourrait dorénavant émerger qui traiterait à la fois des idées et des institutions scientifiques. C'est d'ailleurs l'opinion de David Livingstone (Livingstone, 2003) qui défend que l'application de la posture herméneutique ou « tournant culturel » à l'histoire et la géographie anglaise permet d'appréhender les rapports entre la science et l'espace sous un nouvel angle au début du XXI^e siècle. Avant d'en venir aux implications de ce programme, le prochain chapitre démontre que la volonté croissante en économie et en science régionale et urbaine, depuis les années 1980, de saisir le rapport entre l'espace et l'économie des biens immatériels est également propice au développement d'une géographie des sciences (on pourrait d'ailleurs parler de « tournant culturel » en économie) (Chapitre 2). Cette évolution, qui permet d'envisager les produits scientifiques comme porteurs d'une valeur intrinsèque, rendrait plus accep-

table l'idée d'une géographie de la production scientifique (Chapitre 3). Si l'approche économique des sciences a évolué donnant à la géographie des raisons d'étudier la spatialisation des activités de production scientifique, il conviendrait selon nous, dans la continuité des propositions de David Livingstone (Chapitre 4), de ne pas s'en tenir à une approche « économiste » de l'activité scientifique mais d'intégrer, dans la géographie des sciences, la variété des objets et questions issus des *science studies*.

2.2. *Position sociale, position spatiale et réflexivité*

Jusqu'ici, les évolutions qui ont été évoquées, l'ont été à grands traits, ce qui peut, à tort, donner l'illusion qu'il existe un consensus général sur ces histoires disciplinaires. En réalité, il existe une grande pluralité d'approches, de méthodes et de faits légitimes en sciences humaines et sociales. Dans cette nébuleuse, les cadres disciplinaires servent d'appuis ou de guide pour le positionnement individuel des chercheurs. Ils offrent une matrice stable sur laquelle il est possible de se greffer. Dans une certaine mesure, ils s'ajustent ou plutôt sont ajustables pour convenir à ceux qui s'en réclament. Il faut préciser qu'il ne s'agit pas uniquement d'une affaire de sensibilité. Les disciplines ont des fondements institutionnels plus contraignants, par exemple, que ceux des spécialités scientifiques. Pour Wallerstein, cette organisation, qu'il récuse, résulte de l'adoption d'une option philosophique parmi d'autres, l'option « séculariste ». Pour lui, « la sectorialisation, c'est la présomption que les sciences sociales se divisent en un certain nombre de "disciplines" dont chacune correspond à un discours distinct, intellectuellement défendable. » (Wallerstein, 1982). L'appartenance à une discipline est actée : il existe des coûts à l'entrée puis des responsabilités. Les disciplines ont prétention à faire profession (Ben-David, 1991). Par conséquent, l'appartenance officielle à plusieurs disciplines est une situation qui doit rester exceptionnelle alors que l'appartenance à plusieurs spécialités, bien plus informelle, est tout à fait acceptée². Une bonne façon de connaître et de se représenter une discipline reste donc de recenser ses membres et les revues dans lesquelles ils se reconnaissent (pour l'exemple de la sociologie française, voir Jeannin, 2003). Si cet exercice est envisageable au niveau local voire national, il est plus délicat à mettre en œuvre au niveau mondial et, sans doute, imparfaitement réalisable.

Pour appréhender les champs scientifiques étrangers, il est conseillé d'avoir recours à la littérature : c'est d'ailleurs l'une des missions que doivent remplir les revues de

² À ce sujet, Jean-François Bert évoquait, lors d'une présentation, l'apport et la personnalité de Michel Foucault comme des exceptions en sciences sociales. Jean-François Bert, « Les usages de Michel Foucault dans les sciences sociales », le 3 février 2014, séminaire l'Espace des Sciences Sociales, Labex SMS, Toulouse.

littérature ou les synthèses épistémologiques. Cependant, il n'est pas toujours évident de situer les travaux de synthèse, de savoir quel crédit leur accorder en l'absence de familiarité avec leur contexte de production. Comme Pierre Bourdieu le fait remarquer lors d'une conférence donnée en 1989 à propos des « Conditions de la circulation internationale des idées » : « Les textes circulent sans leur contexte » ce qui génère des « malentendus » qui, d'après lui, pourraient être limités par une meilleure connaissance des « champs de production nationaux et des catégories nationales de pensée. » (Bourdieu, 2002a). Évidemment, Bourdieu met en garde contre ceux qui orchestrent les transferts (les passeurs) et qui ont le pouvoir, aux niveaux nationaux, de donner du poids à leurs interprétations personnelles. Le texte de Bourdieu est intéressant car il suggère – « s'amusant à tenir une position scientifique à l'heure où l'on est postmoderne [les années 1980] » – qu'une bonne connaissance du contexte d'origine d'un document peut permettre d'en faire une lecture plus juste, voire objective. Connaître le contexte d'origine d'un document, c'est aussi pouvoir déceler les mauvaises interprétations et les instrumentalisations, ce qui permet de ne pas être prisonnier des discours « dominants ».

Ce texte de Bourdieu est intéressant à plusieurs égards :

- D'une part, parce qu'en soulignant l'intérêt d'une acquisition de connaissances sur les « champs scientifiques nationaux », il est en mesure de légitimer le développement d'une géographie des sciences. Il est en cela comparable à l'article d'Ophir et Shapin qui lui est contemporain mais qui défend, de son côté, un programme « localiste » projetant d'analyser « les fondements d'une crédibilité ou d'un régime de vérité situés ou localisés » (Ophir & Shapin, 1991).

- D'autre part, cet article témoigne du fait qu'il n'y a pas de consensus sur le statut à accorder aux connaissances, y compris à l'intérieur d'une discipline. Bourdieu se déclare structuraliste et scientifique quand la majorité ou, plutôt, la majorité de ses contemporains anglo-américains est postmoderne et poststructuraliste. Cela se traduit par le fait qu'il veuille établir une hiérarchie entre les interprétations possibles d'un texte. Une posture relativiste voudrait sans doute qu'il n'existe pas de bonnes ou de mauvaises interprétations alors que, pour Bourdieu, les interprétations d'un texte qui ne tiennent pas compte de son contexte d'origine sont de « mauvaises » interprétations. Il se dit « scientifique » parce qu'il croit au progrès scientifique qui n'est possible que s'il existe une hiérarchie entre les idées.

- Pour finir, à l'amorce de l'état de la question, ce texte de Bourdieu est utile pour assoir notre conception des disciplines. On y trouve des arguments capables d'expliquer que l'on puisse distinguer plusieurs « sous-champs » au sein de la géographie et, en particulier, une géographie francophone qui se distingue dans ses préoccupations et ses consensus de la géographie anglo-américaine. Pour autant, les échanges et circulations entre ces différents espaces permettent de maintenir et garantir l'unité de la discipline. À notre

sens, ce qui importe et donne de la consistance aux disciplines à l'échelle mondiale, c'est qu'elles sont le prétexte à un langage commun, à des questions et méthodes de travail communes ou encore à des événements et rencontres partagés. En géographie, cela se traduit par l'existence d'associations comme l'Union Géographique Internationale (UGI) (Plet, 1997), de revues à comités de lecture internationaux et d'événements qui se veulent ouverts au plus grand nombre comme la rencontre Annuelle des Géographes Américains (AAG) ou le Festival International de Géographie de Saint-Dié-des-Vosges (FIG).

Bourdieu suggère que dans l'avancée des connaissances, il y a des tendances dominantes, en particulier aux niveaux nationaux qui, pour des raisons institutionnelles, sont des niveaux d'organisation structurants. Ce sont ces grandes tendances qu'il convient de cerner pour retracer, par exemple, la trajectoire d'un objet d'étude dans une discipline. Cet exercice auquel nous allons nous prêter pour faire l'état de la littérature est difficile car il ne doit pas faire oublier que les personnes et les groupes ont une marge de manœuvre par rapport à la tendance dominante. C'est parce que les traditions scientifiques se perpétuent à travers des « chaînes relationnelles » (l'expression est empruntée à Michel Grossetti) que, par exemple, la tradition quantitative dans la géographie européenne se retrouve inégalement répartie à l'échelle du continent (Cuyala, 2013). De façon analogue, la tendance qui traverse la géographie économique et que Barnes appelle son « tournant culturel » prédomine surtout dans le monde anglo-américain et demeure plus discrète dans celui de la géographie française (Barnes, 2001). Puisque ce qui est dominant quelque part est par ailleurs périphérique, un raisonnement à plusieurs niveaux se justifie pour étudier ce type de phénomène.

À présent que ces nuances et ces éléments fondamentaux pour notre problème et notre objet ont été exposés, il est possible d'entrer au cœur de l'état de la question. Afin de donner un ancrage à ce travail, il convient de revenir sur le traitement de l'objet « science » en géographie. Entre autres intérêts, le préalable qui vient d'être fait permet de signifier que nous avons conscience du caractère structurant de notre position géographique, historique, sociale et institutionnelle. En accord avec ce qui a été établi précédemment, il va de soi que l'ancrage et plus exactement « l'univers de référence » (l'expression est empruntée à Béatrice Milard) auquel se rattache cette thèse dépend largement de son contexte d'élaboration. En particulier, elle s'inscrit dans le champ de la géographie française et s'est développée à partir du programme de recherche interdisciplinaire « Géoscience » portant sur la géographie des activités scientifiques en France et dans le monde. Comme les membres de ce programme étaient historiens, anthropologues, sociologues et géographes, la variété des influences disciplinaires qui nourrissent la présente thèse est significative. Après avoir retracé la trajectoire de l'objet « science » en géographie au cours du XX^e siècle, nous montrerons comment les autres disciplines humaines

et sociales s'en sont emparées avant de revenir sur le parti pris du programme de recherche « Géoscience » et d'étudier le statut de la source bibliographique privilégiée pour les recherches menées dans cette thèse : le *Web of Science*. En particulier, il conviendra d'étudier l'origine de cette source et d'interroger sa pertinence pour la géographie. Compte tenu de notre positionnement, l'état de la question mettra plus particulièrement en avant le contexte de la recherche française. Les travaux étrangers ne seront pas pour autant ignorés puisque le champ scientifique français est soumis à l'influence de nombreux chercheurs et idées qui en sont issus.

CHAPITRE 2. PETITE HISTOIRE DE LA SCIENCE EN GÉOGRAPHIE

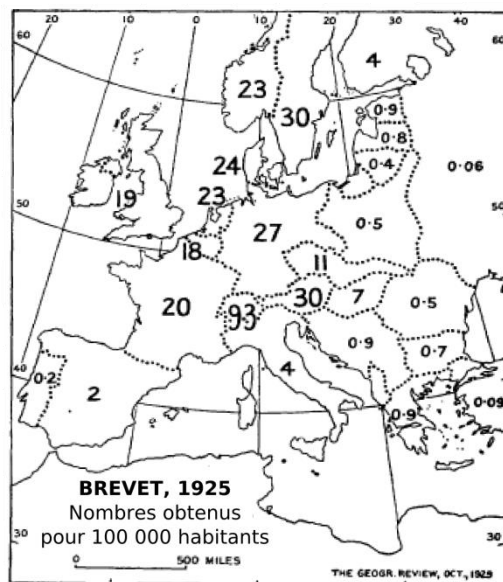
Au cours du XX^e siècle, il n'y a pas de branche de la géographie consacrée à l'objet « science ». En partant de ce constat, il serait acceptable de se contenter d'une revue de littérature tournée vers les disciplines qui s'en sont emparées plus tôt comme la philosophie, la sociologie, l'anthropologie et dans une certaine mesure, l'histoire. Cependant, les activités scientifiques étant une composante difficilement contournable des activités humaines, ce serait une méprise d'affirmer que ces activités ont été parfaitement ignorées par la discipline jusqu'à la fin du siècle. Bien qu'elles n'aient pas bénéficié d'un traitement spécifique, ces activités ont été évoquées dans les travaux d'un petit nombre de géographes tout au long du XX^e siècle. Ceci dit, notre impression est qu'elles ont surtout été abordées en relation avec la dimension économique de la vie sociale. Revenir dans le détail sur la manière dont elles l'ont été et les obstacles qui se sont présentés pour qu'elles bénéficient d'un autre traitement est non seulement envisageable, mais permet en plus de réfléchir aux fondements d'une approche géographique de l'objet « science ». Se prêter à cet exercice est aussi un moyen de s'affranchir d'une critique formulée en 1999 par Marie-Claire Robic, à propos de l'approche géographique de l'objet « mondialisation » : « L'aile marchante de la géographie française a trop tendance à nier ce qui l'a précédée. » (Anthéaume *et al.*, 1999). Pour échapper à une telle limite, il faut avoir conscience que les terminologies sont moins stables que les phénomènes étudiés. Avant de parler de « mondialisation », les géographes s'interrogeaient déjà sur le processus de densification des échanges entre les différentes parties du monde (Arrault, 2007) ; de même qu'avant de parler de « science », les géographes étaient déjà sensibles à la localisation des activités intellectuelles et à la circulation des idées. En fait, les évolutions de la société et le « style de pensée » qui les accompagnent influencent le vocabulaire et les préoccupations des chercheurs. Il existe aussi des contingences, tels des obstacles humains ou techniques qui ont pu empêcher ou retarder le développement de cette thématique de recherche. En particulier, à mesure que le sens associé au terme d'« innovation » a évolué et que les attentes politiques en matière scientifique se sont précisées, il nous semble que le contexte est devenu plus favorable à une géographie de la recherche scientifique (ou géographie des activités scientifiques), et au-delà, si l'on accepte de considérer la science comme un fait social autonome et d'y intégrer les « actes cognitifs », à une géographie des sciences.

1. Idéologie dominante et obstacles pour aborder la science

« La science faite n'a pas de patrie parce qu'elle est impersonnelle. Les Fois, les systèmes en ont parce qu'ils sont toujours personnels et ne peuvent satisfaire que des groupes d'hommes. » (Strada, 1899, p. 160)

1.1. La mesure de la science en géographie

Tout au long du XX^e siècle, la présence de chercheurs et d'activités scientifiques n'est envisagée par les géographes qu'en tant qu'indicateur d'un potentiel dont bénéficie l'espace considéré. Autrement dit, la science n'est pas étudiée pour elle-même par les géographes. Le plus ancien exemple que nous avons identifié est un article signé par Mark Jefferson en 1929 dans la *Geographical Review* et intitulé: « The Geographic Distribution of Inventiveness ». L'ambition de cet article est de cartographier l'inventivité en s'appuyant sur un relevé du nombre de brevets déposés dans plusieurs pays en 1925 (Jefferson, 1929).



Le nombre de brevets obtenus à l'étranger en 1925 par les résidents des pays européens pour cent mille habitants

Figure 1 — Le nombre de brevets obtenus à l'étranger en 1925 par les résidents des pays européens pour cent mille habitants. Carte extraite et traduite de l'article « The Geographic Distribution of Inventiveness » (Jefferson, 1929)

Ce qui intéresse le géographe n'est pas tant la répartition spatiale des inventeurs et des inventions que l'expansion du « monde civilisé ». Son article doit contribuer à cartographier le « monde civilisé » en estimant « la place relative des pays du monde en matière de civilisation » (*ibid.*). Jefferson le publie un an après « The Civilizing Rails », qui traitait de la répartition mondiale des voies de chemin de fer (Jefferson, 1928). Des deux articles, « *The Civilizing Rails* » est devenu le plus connu : il a été deux fois plus cité que « *The Geographic Distribution of Inventiveness* » (43 citations contre 14 à l'été 2015 dans *Google Scholar*). Cette différence ne surprend pas car, entre-temps, la géographie des transports s'est développée pour figurer parmi les spécialités les plus prolifiques de la discipline. Si l'on se fie à *Google Scholar*, la plupart des chercheurs qui ont cité « *The Geographic Distribution of Inventiveness* » s'inscrivent non pas dans le champ de la géographie, mais dans celui des sciences de l'innovation, comme par exemple Jeremy Howells et Gaétan de Rassenfosse. Depuis 1990, l'article de Jefferson a tout de même été cité par un géographe, Yves Guermond, professeur à l'Université de Rouen. Dans « Y-a-t-il des territoires innovants ? » Guermond mobilise cette citation pour revenir sur le traitement géographique de l'objet « innovation » (Guermond, 2001). Que Guermond se penche sur le traitement de l'innovation plutôt que sur celui de la science est symptomatique. Jusqu'à la fin du XX^e siècle, la science, en tant que domaine d'activité institué, autrement dit « la recherche scientifique », n'est presque jamais abordée directement par les géographes. Après avoir été associée à la civilisation puis au développement, elle est plus concrètement envisagée à travers son lien à l'innovation, ou exceptionnellement à la formation à partir des années 1970-1980. Comme nous allons le voir, le fonctionnement interne de la science n'intéresse longtemps qu'à la marge l'économie, les sciences politiques, statistiques et administratives, autant d'univers cognitifs dans lesquels baigne la géographie à l'heure où celle-ci développe sa dimension appliquée.

Dans son article, Guermond insiste sur les difficultés rencontrées par Jefferson pour manier et interpréter les chiffres de dépôts de brevet. Le traitement de ce type de données pose des problèmes multiples de mise en forme, comptage, normalisation et significativité. En son temps, Jefferson reçoit plusieurs critiques concernant la pertinence de ses traitements statistiques, la plus notable étant celle de Seabury C. GilFillan, un sociologue de Chicago favorable au développement d'une sociologie de l'invention (GilFillan, 1930). Compte tenu de ces spécificités, la mesure de « tout ce qui est en lien avec l'activité scientifique » a depuis justifié l'existence de la scientométrie, une spécialité autonome à partir des années 1970, sur laquelle nous reviendrons (Polanco, 1995) ; et comme Jérôme Lamy le souligne pour expliquer la réticence de certains historiens à faire de l'histoire des sciences, la particularité des sources de la scientométrie a sans doute rendu difficile la prise en main de l'objet « science » et du reste la poursuite d'une approche

empirique des activités scientifiques (Lamy, 2003). Tout comme les historiens de la Nouvelle Histoire, les géographes n'ont pas l'habitude des données sur la science. Contrairement aux données sur les infrastructures de transport, le recensement, les navettes domicile-travail ou encore le commerce international, les données portant sur l'activité scientifique ne sont pratiquement jamais exploitées en géographie.

Après Jefferson, un autre géographe américain a toutefois manipulé ce type de données, c'est Stephen Sargent Visser (1887-1967), mieux connu pour ses travaux de climatologie. À la suite d'un pionnier de la scientométrie, le psychologue James McKeen Cattell, Visser produit des analyses statistiques spatialisées sur les scientifiques américains pendant 25 ans (1922-1948), non sans un certain penchant déterministe, voire « eugéniste » (Godin, 2006b, p. 46 ; Visser, 1948).

Brillants scientifiques nés en Indiana							
Nombre de naissances par science et décennie de naissance							
Né..... avant 1860	1860's	1870's	1880's	1890's	1900's	1910's	Total
Anatomie.....	0	0	1	0	1	0	2
Anthropologie...	0	1	1	0	0	0	2
Astronomie.....	2	0	1	2	0	0	5
Botanique.....	1	5	3	2	1	0	12
Chimie.....	2	0	2	1	3	3	11
Géologie.....	1	3	1	0	0	1	6
Mathématiques.	0	2	2	0	0	0	4
Pathologie.....	1	1	0	1	0	1	4
Physique.....	1	2	0	0	2	2	7
Physiologie.....	0	1	1	1	0	0	3
Psychologie.....	0	3	1	1	1	0	6
Zoologie.....	3	3	7	5	0	1	19
—	—	—	—	—	—	—	—
Totaux.....	11	21	20	13	8	8	81
Population							
à mi-année							
Millions.....	1.2	1.5	1.8	2.1	2.35	2.6	
Profit par							
Millions.....	9.1	10.0	11.0	6.2	3.4	3.1	0

Tableau 1 — Brillants scientifiques nés en Indiana. Extrait et traduit de l'article « Starred Scientists Born or Trained in Indiana » (Visser, 1948)

Puisque Visser effectue ce travail à côté de ses recherches en géographie physique, il ne développe pas de programme de recherche spécifique sur cette question et ne transmet pas son savoir-faire aux générations suivantes de géographes. Au bout du compte, en dehors de quelques exceptions ponctuelles au rang desquels se placent les traitements statistiques et cartographiques de Madeleine Brocard sur la géographie de la recherche française (Brocard, 1978, 1981, 1986, 1991 ; Brocard, Hérin, & Joly, 1996), les géographes du XX^e siècle ne peuvent être crédités d'aucune expertise en matière de traitement de données scientométriques.

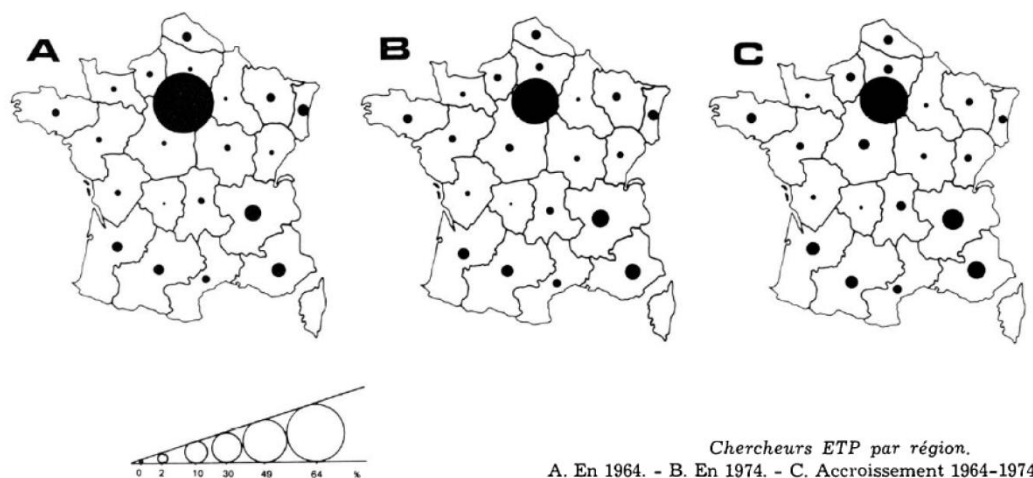
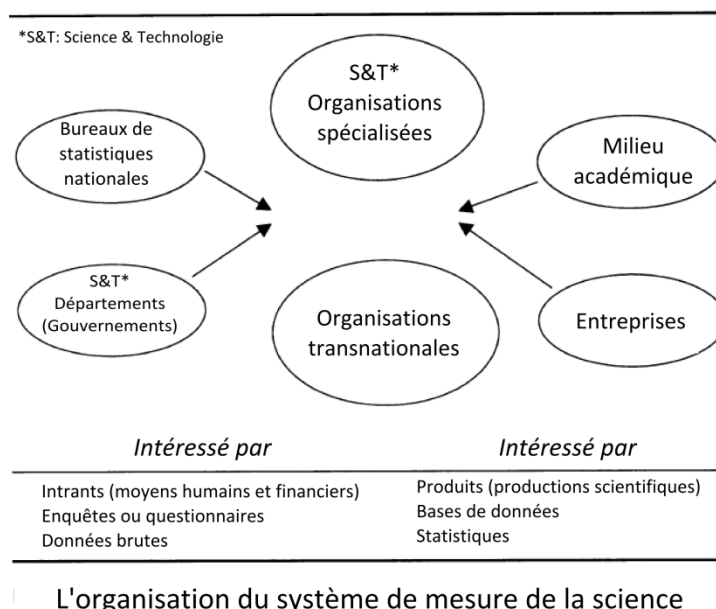


Figure 2 – Chercheurs ETP (Équivalent temps plein) par région. Carte extraite de l'article « Aménagement du territoire et développement régional : le cas de la recherche scientifique » (Brocard, 1981)

Pour comprendre le peu d'attention des géographes pour ces données, il est essentiel de noter que les organismes de la statistique publique n'ont eu recours que tardivement aux indicateurs scientométriques. En effet, Brocard souligne à plusieurs reprises le manque de données administratives disponibles pour étudier la régionalisation de la recherche française (*ibid.*, 1991, p. 46-49, p. 220-226). Une anecdote permet d'élargir ce constat au niveau mondial. Il s'agit d'une plaisanterie de Stevan Dedijer dans son plaidoyer pour la diffusion d'une science de la science : « d'aucuns se rappellent peut-être de ce qui se disait en 1960 à propos de 147 des 150 pays du monde : ils connaissent le nombre de leurs ânes mais pas le nombre de leurs chercheurs. »³ (Dedijer, 1966). À partir des années 1950, quelques pays récoltent des données sur l'investissement économique et les ressources humaines dans la recherche et le développement (*input*), mais ce n'est qu'à partir des années 1980 qu'ils commencent prudemment à se pencher sur la performance de leurs scientifiques (*output*). C'est ce que montre Benoît Godin constatant que les premiers utilisateurs des chiffres portant sur la quantité et la visibilité des publications scientifiques (*output*) sont les chercheurs eux-mêmes (Godin, 2002). D'après Godin, mais c'est aussi l'avis du spécialiste des relations scientifiques internationales Aant Elzinga, si les organismes de statistiques gouvernementales se focalisent essentiellement sur les mesures d'*input* jusqu'aux années 1980, c'est en raison de l'influence de l'Organisation de Coopération et de Développement économique (OCDE) qui inspire à ses pays membres leurs méthodes de mesure de la science et leurs politiques scientifiques (Elzinga, 2001, 2012). Les re-

³ L'original : « one may recall what was said in 1960 of 147 of the world's 150 countries: they know the number of their donkeys, but not the number of their research workers. » (Dedijer, 1966).

commandations de l'OCDE pour la mesure des intrants de la recherche sont d'ailleurs régulièrement consignées depuis 1963 dans un document de référence appelé le *Manuel de Frascati* (OCDE, 1963).



L'organisation du système de mesure de la science

Figure 3 – L'organisation du système de mesure de la science. Schéma extrait et traduit de l'article « *Outline for a History of Science Measurement* » (Godin, 2002)

1.2. L'approche de l'OCDE : enjeux économiques et politiques

L'approche de l'OCDE s'inspire de la logique de la *National Science Foundation* (NSF) elle-même conçue en accord avec les positions de Vannevar Bush. La vision de cet ingénieur et conseiller présidentiel américain, connu pour avoir occupé un poste clef à l'Office de la Recherche Scientifique et du Développement (OSRD) pendant la Seconde Guerre mondiale, est résumée dans un texte célèbre, *Science, the Endless Frontier* (Bush, 1960). Sa thèse est qu'au lieu de guider ou contrôler les chercheurs, les politiques scientifiques doivent supporter la recherche fondamentale. D'après ce rapport, les États ont deux bonnes raisons de développer une politique scientifique et de soutenir leurs chercheurs, d'une part pour le développement économique de leur pays et d'autre part, pour la paix entre les nations.

Godin remarque que pour persuader les gouvernements de financer la recherche et protéger l'autonomie des scientifiques, le rôle des économistes a été crucial. Par exemple, dans les années 1950-1960, Richard R. Nelson et Kenneth Arrow soutiennent l'idée que la « science est un bien public pur, ce qui veut dire que les bénéfices qui en résultent, ne

sont pas appropriables par leurs producteurs. Ils circulent librement et génèrent des résultats et des impacts difficiles à mesurer et à contrôler. »⁴ (Godin, 2002). Ces arguments font autorité jusqu'aux années 1970 pour laisser la science à l'état de « boîte noire ». En plus, suivant le schéma linéaire attribué à Joseph Schumpeter, la recherche fondamentale précède la recherche appliquée qui précède l'innovation, et donc le développement socio-économique (voir la Figure 4, issue de Noailles, 2011). Il est également courant, bien qu'évidemment réducteur, d'attribuer à Schumpeter la valorisation du concept d'« innovation » au détriment de celui d'« invention ». Ce faisant, sa théorie aurait favorisé l'association de « science » et « innovation », deux concepts qui n'avaient rien à voir ensemble jusqu'au XIX^e siècle (Godin, 2014). Dans cette logique, les investissements en matière scientifique se traduisent mécaniquement par de l'innovation si bien qu'il serait vain de les cibler, ou de chercher à en mesurer les débouchés immédiats sous forme de publications ou de brevets pour en contrôler l'efficacité. Typique de cette vision, un article d'Ira Horowitz paru en 1966 démontre l'existence aux États-Unis d'une coïncidence directe entre la présence des scientifiques dans une région (*input*) et la croissance économique de cette région (*output*) (Horowitz, 1966).

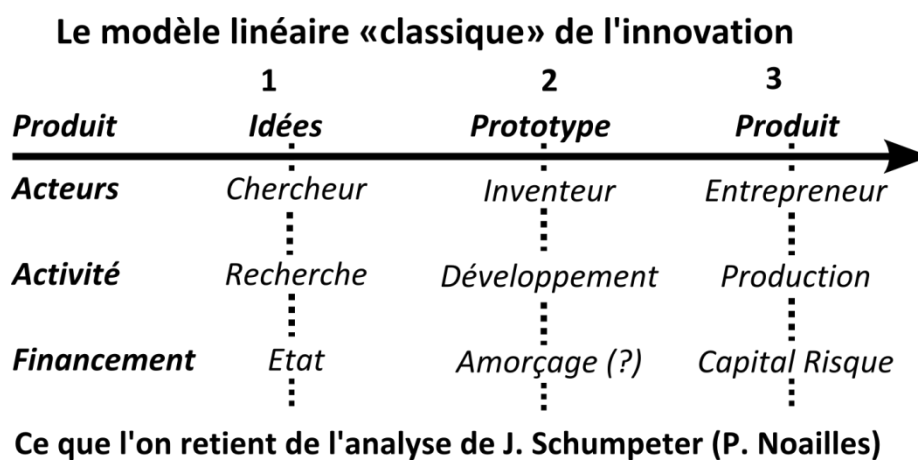


Figure 4 – Le modèle linéaire « classique » de l'innovation. Reproduction du schéma de l'article « De l'innovation à l'innovateur : Pour une approche structuraliste de l'innovation » (Noailles, 2011)

⁴ L'original : « The economist views science as a pure public good : science gives rise to benefits that, contrary to private goods, are not appropriable by the producer, and it generates results and impacts that are difficult to measure and control (Nelson 1959 ; Arrow 1962). » (Godin, 2002).

Cette idée d'une science génératrice de croissance est soumise à rude épreuve au moment des crises des années 1970, tandis que la théorie de la croissance inspirée des travaux de l'économiste Robert Solow, dans laquelle le progrès technique est une variable exogène, est en passe d'être révisée (voir Section 2). Progressivement, le lien direct entre science et innovation cesse d'être pris comme une évidence. Les résultats de l'économiste Zvi Griliches, sur lesquels nous reviendrons (dans la Section 2 de ce chapitre), nuancent l'impact direct de la science sur la croissance et laissent penser qu'il serait possible et plus pertinent de mesurer les effets de la recherche, non pas sur la croissance, mais sur l'« innovation », qui peut se mesurer, par exemple, par le nombre de brevets déposés (Griliches, 1979 ; Godin, 2002). Compte tenu de la détérioration de la conjoncture, les politiques scientifiques s'appuyant sur le modèle linéaire ne sont plus tenables. Les recherches les plus susceptibles de favoriser l'innovation, c'est-à-dire capables d'avoir des retombées économiques, entrent dans les bonnes grâces des politiques publiques (Gingras, 2003). À l'heure d'une réduction des dépenses militaires, il s'agit aussi d'inciter le secteur privé à investir dans la recherche, notamment biomédicale, même si l'État continue de jouer un rôle central pour l'économie (Chomsky, 2009). Ce changement d'attitude par rapport à la production de connaissances s'explique par les restrictions budgétaires des gouvernements mais aussi par la nécessité de répartir les moyens entre davantage de centres d'activité scientifique et de champs disciplinaires : c'est la fin de « l'État providence ». Il en résulte un intérêt croissant des organismes officiels pour les indicateurs scientométriques qui portent sur la production scientifique et son impact (*output*). Cette évolution va de pair avec la valorisation dans les discours officiels des impératifs de compétitivité et d'excellence au détriment de la focalisation sur le progrès et le développement.

En plus d'une vision économique, l'OCDE relaye une vision internationaliste de la science : la science est par essence internationale dans la mesure où les vérités scientifiques se doivent d'être universelles (Godin, *op. cit.*). Puisque la science fructifie grâce aux coopérations internationales elle est, au même titre que le commerce, en position de favoriser l'entretien de relations pacifiques entre les États. Cette idéologie qui porte le nom d'« internationalisme scientifique » est largement mobilisée par les États-Unis au sortir de la Seconde Guerre mondiale (Elzinga, 2012 ; Miller, 2006). Elle est au cœur du célèbre discours du Président Eisenhower, prononcé le 8 décembre 1953 à l'Assemblée générale des Nations Unies, et intitulé « *Atoms for Peace* » (Eisenhower, 2003). Dans le cadre de l'OCDE, cette idéologie légitime la mise en place d'aides à la coopération scientifique internationale qui sont, pour partie, des aides au développement. Ces aides au développement ont clairement une dimension géopolitique et sont, au temps de la guerre froide, un moyen de s'assurer l'alliance des pays du Tiers-Monde. Dans ce contexte postcolonial, il n'est plus question d'apporter unilatéralement « la civilisation » à des peuples colonisés

mais d'aider des peuples libres ou en train de prendre leur indépendance à se « développer » (Wallerstein, 1982). Cette approche solidaire et bienveillante se transforme également à partir des années 1970. D'après Armand Mattelart, « la conception linéaire du développement/progrès, inspirée de l'histoire industrielle de l'Occident » entre en crise dans les années 1970 et les pays occidentaux commencent à laisser de côté « la question de la gestion internationale de l'inégalité de développement » (Mattelart, 1999). Les formes d'aides à la coopération scientifique internationale évoluent, d'après Philippe Losego et Rigas Arvanitis, pour prendre une tournure contestable : « Disposant d'instruments coûteux, élaborés en collaboration avec des entreprises sises sur leurs territoires et aidés par des programmes scientifiques proches de leurs thématiques, les laboratoires des pays hégémoniques tendent à transformer les laboratoires des pays non hégémoniques en "succursales" de la recherche mondiale. » (Losego & Arvanitis, 2008). D'une manière plus générale, l'optimisme associé à l'idée de progrès diminue car « pour la première fois, la productivité liée aux avancées technologiques supprime des emplois sans contribuer à des créations dans d'autres domaines, en nombre égal ou supérieur aux emplois supprimés, d'où une réduction du volume global de travail » (Dollfus, 1997, p. 166). L'apparition d'un chômage de masse dans les « pays développés » s'accompagne d'une inquiétude quant aux impacts négatifs du progrès technique sur la planète. Il s'en suit une modification des attentes envers les scientifiques et une volonté plus forte de contrôler les avancées et les orientations des recherches.

En fait, à plusieurs niveaux institutionnels, un désengagement financier des pouvoirs publics justifie l'adoption d'une doctrine « néolibérale » qui se traduit par la nécessité de privilégier des financements ciblés (Paradeise, Reale, & Goastellec, 2009). Les dépenses investies dans la recherche doivent rapporter et le progrès doit aller dans le bon sens. Les effets de la transition démographique et les préoccupations environnementales sont, par exemple, mis à l'ordre du jour des agendas de recherche internationaux comme Europe 2020. Idéologiquement, l'heure n'est plus à la solidarité mais à la compétition. La stratégie de Lisbonne, mise en place au Conseil Européen de mars 2000, et dont il est convenu d'admettre qu'elle s'est soldée par un échec (Cohen-Tanugi, 2009), se veut emblématique de la volonté de développer l'« économie de la connaissance la plus compétitive et la plus dynamique du monde » dans le cadre d'un Espace Européen de la Recherche (EER) qui reste encore à consolider (Garrigue, 2003, p. 55). Face à un nombre plus important de sites, d'organisations et d'humains impliqués dans la recherche et en raison d'une situation financière délicate pour de nombreux États, les solutions retenues s'inspirent du « Nouveau Management Public » (NPM) – mode de gouvernance résultant de l'application aux secteurs publics des techniques d'administration et de gestion des grandes entreprises – pour ne pas gaspiller d'argent, il faut cibler les investissements en

amont et trouver un moyen d'en mesurer les résultats en aval. Les universités ont un rôle central à jouer dans ce processus, et cela dans plusieurs pays européens (Ferlie, Musselin, & Andresani, 2009). C'est donc parce que premièrement la part des fonctionnaires s'adonnant à la recherche devient considérable, et deuxièmement parce que le coût humain et matériel de certaines pratiques scientifiques se renforce dans un contexte où les ressources publiques s'amenuisent, que l'idée de contrôler l'efficacité des unités de recherche publiques se met en place et au-delà, l'idée d'évaluer les établissements auxquels elles se rattachent (Arvanitis, 1983).

Ces évolutions expliquent qu'en matière de politique scientifique, deux grandes phases sont généralement distinguées : celle qui s'étend jusqu'à la fin de la guerre froide et s'accompagne d'une conjoncture économique favorable dans les pays industrialisés avec une croyance dans le progrès et le développement ; et celle qui démarre dans les années 1980. Marquée par une conjoncture économique maussade et la désindustrialisation dans les pays du Nord, cette seconde phase est souvent présentée comme propice à la suprématie de la recherche appliquée et privée. On retrouve cette vision linéaire et dichotomique dans la théorie de Michael Gibbons *et al.*, nous y reviendrons, qui distingue deux modes de production scientifique : le premier qui valorise la recherche fondamentale et l'autonomie de la recherche (mode 1) et le second qui se focalise sur la recherche appliquée et l'établissement de liens entre science et innovation, autrement dit secteurs publics et privés (mode 2) (Gibbons, Limoges, Nowotny, & Schwartzman, 1994). Dans la suite de cet argumentaire, tout en sachant que la réalité est loin d'être aussi binaire et qu'il y a, au contraire, une co-existence des deux systèmes, nous montrons que la façon dont les activités scientifiques ont été envisagées par les géographes à partir du milieu du XX^e siècle fait écho à l'approche de la science véhiculée par les politiques publiques inspirées des préconisations de l'OCDE.

1.3. La science : ressource ou richesse ?

En géographie, on distingue deux grandes périodes. La première période (1950-1970) est une période au cours de laquelle l'attention des géographes porte sur le développement : ce n'est pas la science elle-même qui compte mais la diffusion d'un progrès essentiellement technique et médical. Pendant la seconde période (1980-2000), la science est envisagée à travers les activités de recherche qui sont perçues comme vecteur de compétitivité pour les territoires, y compris aux niveaux régional et urbain où elles accompagnent le processus de décentralisation politique. En dépit de cette évolution, les activités scientifiques n'intéressent qu'à travers leur rôle pour la société ou le monde « extérieur », et cela jusqu'à la fin des années 1990. Cette appréhension se traduit par le recours à la

terminologie « recherche » au détriment de « science ». En effet, d'après la définition de Madeleine Brocard inspirée du manuel de l'OCDE : « les activités de recherche et de développement, et les activités scientifiques extérieures à la recherche se distinguent essentiellement par la présence ou l'absence d'un élément de nouveauté et d'innovation : lorsqu'une activité s'exerce selon une routine établie, elle n'appartient pas à la recherche et au développement ; lorsqu'elle s'écarte de la routine et fraye une voie nouvelle, elle peut être considérée comme "recherche et développement". » (Brocard, 1991, p. 262). Opérer cette distinction, qui mériterait d'être discutée, et insister sur la recherche au détriment de la science, permet de mettre l'accent, implicitement, sur la portion de l'activité scientifique qui peut se traduire par des retombées économiques.

En fin de compte, il faut attendre la transformation manifeste au cours des années 1990 de la valeur économique associée à l'information et à la connaissance pour qu'une autre approche de la science devienne possible en géographie. Petit à petit, les économistes en viennent effectivement à théoriser les principes d'une « nouvelle économie » (Stiglitz & Walsh, 1993), une économie « immatérielle » (Bomsel, 2010), ou de la « connaissance » (Foray, 2000). Dans ce contexte, Partha Dasgupta et Paul A. David parlent même d'une « nouvelle économie des sciences » (Dasgupta & David, 1994), mais Michel Callon et Dominique Foray objectent qu'il s'agit plus précisément de la montée en force d'une socio-économie de la recherche scientifique qui s'accompagne d'une économie industrielle de la science (Callon & Foray, 1997). Si elle donne lieu à de nouveaux courants de recherche dans les années 1990, l'idée que le savoir est richesse (en soi) est pourtant loin d'être neuve ; en témoigne le surnom que donne l'illustre Francis Bacon à une partie des savants de *La Nouvelle Atlantide* : « les marchands de lumière » qui sont, dans cette utopie publiée en 1627, les sages chargés de sillonner les mers à la recherche de nouvelles connaissances et d'inventions à rapporter sur l'île de Bensalem (Bacon, 1627). En fait, Bacon a été parmi les premiers à décrire la connaissance scientifique comme la plus précieuse des ressources dont un pays peut disposer. Par ailleurs, il est l'auteur de référence quand il s'agit de remonter aux origines du système scientifique contemporain (Dedijer, 1966). À l'intendant de Bensalem, Bacon fait dire : « Nous maintenons donc un commerce, non pour nous procurer de l'or, de l'argent, des pierres précieuses, des soieries, des épices ou d'autres commodités matérielles, mais pour la première Œuvre de Dieu qui fut, la Lumière » (Bacon, *op. cit.*, p. 106). La métaphore mercantile peut paraître étonnante car il n'est pas question, pour les « marchands de lumière », de s'enrichir en vendant leurs savoirs ; mais, pour l'utopiste, c'est l'occasion d'évoquer le fait que les connaissances scientifiques peuvent être transportées et échangées.

Comme les filières du diamant ou de la rose sont des objets indiscutables pour les géographes soucieux de mettre en évidence les dynamiques, les enjeux et les rouages du

« système-monde » (R. Brunet, 2003 ; Calas, 2013), l'organisation de la production scientifique au niveau mondial ne semble digne d'intérêt qu'à partir du moment où la connaissance est perçue comme une richesse à part entière plutôt que comme un intermédiaire pour obtenir des richesses. Si cette approche mercantiliste de la science facilite son entrée dans la géographie, un risque demeure à raisonner de la sorte : c'est celui de se limiter à la métaphore marchande pour finalement ne jamais considérer le fait scientifique dans toute sa splendeur (ses différentes dimensions et niveaux d'action). À ce titre, nous défendons l'idée qu'un rapprochement avec l'histoire, l'anthropologie, la sociologie et plus précisément, les *science studies* est primordial pour une géographie des sciences. Ainsi, pour que la science bénéficie d'un traitement adapté à son caractère et à son autonomie vis-à-vis des autres faits sociaux, nous pensons légitime de lui consacrer une branche de recherche spécifique au lieu que, comme c'était le cas tout au long du XX^e siècle, elle continue d'être évoquée à la marge par la géographie économique, la géographie urbaine ou encore la géographie politique (Limonier, 2012 donne un exemple récent de son traitement en géographie politique).

En attendant, la section suivante montre que l'importance accordée aux activités intellectuelles a augmenté en géographie économique et urbaine. L'hypothèse guidant ce chapitre est qu'au moins dans le cas français, c'est plus spécifiquement par le biais de l'innovation que la localisation des activités scientifiques a pris de l'importance. Familière aux géographes, la notion d'« innovation » acquiert une nouvelle consistance en science régionale et urbaine, et au-delà dans les politiques publiques à partir des années 1980. C'est à ce moment-là que la réticence des géographes à s'occuper des biens immatériels commence à s'estomper. Pour en témoigner, le plaidoyer d'Henry Bakis pour une géographie des télécommunications est éloquent (Bakis, 1980). Toutefois, il faut encore patienter une trentaine d'années pour que la géographie se mette sérieusement à envisager la science comme objet de recherche. Entre temps, la recherche scientifique est surtout perçue comme une activité du secteur tertiaire et sa dimension productive reste peu explorée, d'où son inclusion dans le volume « Service et commerce » de l'Atlas de France en 1999 (Saint-Julien, 1999). Même dans la géographie de Roger Brunet, le stock de science est le plus souvent à mettre du côté de l'information, « l'une des grandes énergies des systèmes géographiques, à travers laquelle la détection et l'exploitation des ressources sont possibles » (R. Brunet, 2001, p. 277). En fait, plusieurs manières d'envisager les productions scientifiques peuvent se justifier : comme des ressources, voire des richesses en soi, ou bien comme des moyens de connaître, valoriser et exploiter les ressources et créer les richesses (autrement dit comme instruments de médiation). Jusqu'à la fin des années 1990, c'est la seconde approche qui domine, retardant sans doute le développement d'une géographie des sciences.

2. De la diffusion des innovations à la géographie de l'innovation

« Le Monde est fait d'espaces différents, dont les systèmes n'absorbent pas tout et n'importe quoi. » (R. Brunet, 2001, p. 81)

2.1. De la diffusion des innovations...

Dans la géographie du XX^e siècle, les activités de recherche ont généralement été évoquées implicitement à travers des indicateurs de développement ou assimilées aux activités tertiaires (Rocheft, 1957 ; Polèse, 1974 ; Claval, 1977 ; Brocard, 1978). Les limites et le caractère trop composite de cette dernière catégorie, qui désigne « toutes les professions non productrices » (Rocheft, *ibid.*), sont pourtant connues mais, dans certains cas, jugées acceptables.

Par exemple, en 1957, Michel Rocheft l'utilise pour étudier les « réseaux urbains » et s'en explique : « Malgré l'intérêt que pourrait présenter une analyse de la composition interne de ce secteur [le secteur tertiaire], cette démarche nous semble peu rentable dans le cadre de l'objet précis qui nous occupe. En effet, chaque centre combine toujours plus ou moins tous les "services", et le secteur tertiaire dans sa complexité traduit mieux cette action globale du centre sur sa région. Ajoutons qu'en travaillant sur des nombres plus élevés on élimine plus facilement les facteurs occasionnels. » (Rocheft, *ibid.*). En revanche, pour étudier le cas particulier de la recherche comme entend le faire Madeleine Brocard pour sa thèse d'État, le recours à la catégorie tertiaire est présenté comme insatisfaisant (Brocard, *op. cit.*, p. 2). Et dès le milieu des années 1970, à mesure que la part des emplois concernés par le secteur tertiaire augmente dans la population active (Polèse, *op. cit.*), les critiques sur cette catégorie se multiplient, à commencer par celle qu'Albert Tauveron publie en 1974 dans l'*Espace Géographique* : « cette expression ne rend pas compte clairement de la réalité du développement des services à la production ou à la reproduction sociale, qu'ils soient publics ou privés : la différenciation s'accroît entre les activités de services liées aux phases stratégiques de l'investissement et les activités d'exécution et de mise à disposition. Leurs localisations ont, par conséquent, des rationalités différentes qu'il est nécessaire d'analyser pour rendre compte des aspects apparemment contradictoires de la croissance régionale des services (concentration sur les capitales, décentralisation de grands groupes). » (Tauveron, 1974). Aussi, dans son « Que-Sais-Je ? » sur *la Nouvelle Géographie*, Paul Claval suggère qu'il existe plusieurs types de services que les modèles de communication, qui s'imposent dans les années 1970, permettent mieux

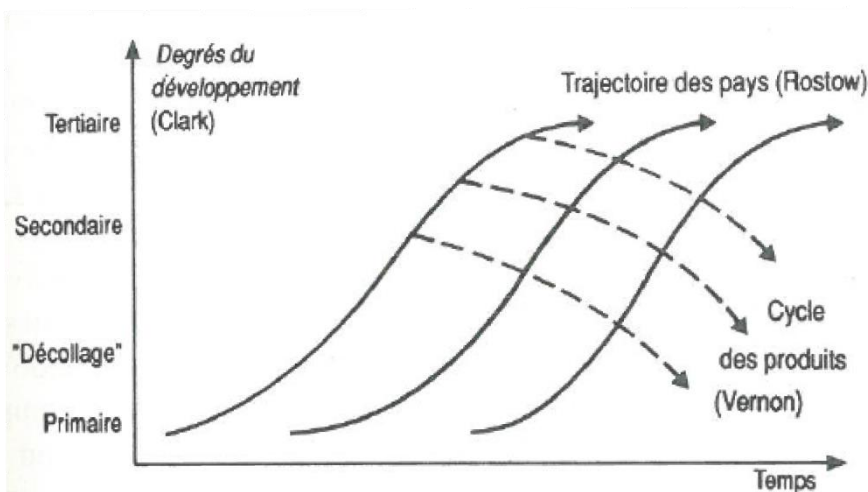
de différencier que la théorie classique des lieux centraux : « [Les modèles de communication] s'appliquent à des secteurs économiques jusque-là négligés. Les services qui doivent traiter très rapidement des informations changeantes – les agences de presse par exemple – sont liés aux nœuds du système de communication, aux grandes métropoles ; les services qui traitent d'une information qui peut se stocker et ne se démode pas aussi vite sont beaucoup plus libres dans leurs choix : les grandes entreprises installent volontiers leurs bureaux de recherche ou leurs banques de données dans des villes moyennes. » (Claval, 1977, p. 89). Enfin, en 1985, Antoine Bailly met l'accent sur les problèmes associés au caractère composite de cette catégorie tertiaire pour étudier le développement régional suisse (Bailly, 1985).

En plus d'envisager les activités de recherche scientifique comme non productrices, les géographes d'après-guerre ont tendance à préférer la notion de « techniques » à la notion de « science ». Pour Jean Gottmann, « l'ensemble des ressources dont dispose un État n'est pas tant déterminé par la texture du sol ou du sous-sol, par le climat ou la géologie, que par l'ensemble des techniques et des moyens de mettre en œuvre les techniques dont il dispose » (Gottmann, 1951, p. 163). Aussi, dans le « Que-Sais-Je » d'Yves Lacoste sur *Les pays sous-développés* (Lacoste, 1959), il est constamment question d'inégalités entre pays du point de vue des compétences techniques, mais jamais de différences d'avancement scientifique. Dès lors, pour se développer, la priorité consiste à imiter ou adopter les techniques venues d'ailleurs, et non à se donner les moyens de mettre au point de nouvelles techniques. D'après la remarque de l'économiste germano-britannique Hans Singer reprise par Lacoste : « Il est beaucoup plus facile de transporter les fruits du développement industriel que d'en transporter les semences » (Singer cité par Lacoste, 1959, p. 54). Cela signifie que le développement exogène est central pour enrayer le « sous-développement ». Si les chercheurs et les enseignants des pays développés doivent jouer un rôle dans ce processus, c'est, par exemple, à travers la diffusion des « progrès de la médecine » : « leur application philanthropique dans de nombreux pays » contribuant à la poussée démographique observée dans les pays sous-développés (*ibid.*, p. 74). En fait, à cette époque-là, le savoir-faire technique et l'accès aux soins sont des variables jugées bien plus importantes dans le rapport des hommes à leurs milieux, et en tant que vecteurs du changement social, que le niveau d'avancement scientifique. L'importance de l'accès au soin contribue même à l'émergence d'une branche de la géographie consacrée à la santé, qui a connu un développement considérable depuis.

À nouveau, l'analogie avec l'histoire fonctionne puisque les historiens français ont longtemps privilégié l'étude historique des techniques et des sciences médicales sur l'étude historique des sciences en général (Lamy, 2003). Au sujet de la santé, Lamy écrit : « La dimension foncièrement sociale de la médecine, cette discipline qui cerne l'homme

dans ses arcanes les plus intimes et qui investit le corps comme territoire d'analyse, intéressait les historiens. En fait ces derniers n'ont pas cherché à exploiter un secteur spécifique des sciences ; ils ont simplement trouvé l'histoire de la santé au cœur de leurs questionnements. L'école des Annales plaçant l'homme au centre de ses préoccupations. » (Lamy, 2003). Tout comme les historiens, les géographes français se sont longtemps focalisés sur les dispositifs scientifiques ayant un impact socio-économique positif et direct sur les sociétés humaines. Ce qui importait pour eux n'était donc pas le progrès technique ou médical en tant que tel, mais le fait que la population puisse accéder aux fruits de ce progrès. L'intérêt manifeste pour la diffusion spatiale des innovations est caractéristique de cette logique.

Dans son introduction à l'ouvrage collectif Les Régions qui gagnent (p. 20-21) George Benko commente le schéma ci-dessous intitulé la "dualité Rostow-Vernon".



"Dans les années 1960 domine une première orthodoxie en matière de développement spatial des activités économiques. Chaque aire géographique (région ou pays) serait censée passer par les mêmes étapes du schéma historique de Colin Clark (1951), les âges pré-industriel (primaire), industriel (secondaire), post-industriel (tertiaire voire quaternaire). Mais tous les régions ou pays ne décolleraient pas au même moment, d'où le développement relatif des uns par rapport aux autres à chaque moment de l'Histoire. Telle fut la théorie des étapes du développement de W. Rostow (1963). Transversalement à ce décalage des aires géographiques, les nouveaux produits inventés dans les zones les plus développées se banaliseraient et leur production se déplacerait vers les contrées moins développées (cycle du produit de Vernon, 1966). Le schéma Clark-Rostow-Vernon n'est pas, selon la classification des approches esquissée plus haut, globalement structuraliste. Rien n'empêche qu'à la fin des temps, dans l'ère quaternaire, les trajectoires de tous les pays convergent vers une structure interne semblable."

Encadré 1 - La dualité Rostow-Vernon. Schéma extrait de l'Introduction à l'ouvrage collectif Les Régions qui gagnent (Benko, p. 20-21 in Benko & Lipietz, 1992)

Pour simplifier, on considérerait que le développement venait de l'extérieur, qu'il était exogène et dépendait localement du rythme d'adoption des innovations, qu'importe leur provenance. Cette idée se nourrissait de la théorie, par rapport à laquelle il fallait se positionner, des stades de développement de Rostow (Encadré 1), paradigme dominant à l'époque (Benko, 2007 ; Benko & Lipietz, 1992).

Pour analyser la diffusion spatiale des innovations, les propositions de Torsten Hägerstrand ont longtemps fait, et font encore, référence. Elles sont à l'origine de ce que Claval appelle les « modèles de communication » (Claval, *op. cit.*). Connue comme le chef de file de « l'École de Lund », Hägerstrand publie en 1953 *Innovation diffusion as a spatial process* – titre de la traduction anglaise réalisée par Allan Pred en 1967 – (Hägerstrand, 1953). En tant que géographe, Hägerstrand s'intéresse à « l'interprétation spatiale des phénomènes sociaux » (*ibid.*). Ce qui est original dans sa géographie, c'est qu'il n'entend pas se limiter aux innovations qui se manifestent physiquement dans le paysage. Il propose une grille d'analyse permettant de considérer toutes les innovations qui, aussitôt qu'elles sont adoptées, deviennent « caractéristiques de leurs régions d'adoption » (*ibid.*). En ce sens, les idées, les modes, les produits culturels (et même les découvertes scientifiques !) ont leur place dans son modèle. Ce modèle est influencé par l'école diffusionniste de Berkeley, et en particulier par le travail de Carl O. Sauer sur la diffusion des techniques agricoles, mais pas seulement. D'après Peter Haggett, il fait la synthèse de trois approches : celle de Sauer, celle de l'historien Frederick Jackson Turner sur le thème de la frontière américaine, et celle des sociologues, comme Everett Rogers, qui s'intéressent à la propagation des idées (Haggett cité par Gondard, 1991). Ainsi, Hägerstrand parvient à mettre au point des méthodes quantitatives de modélisation valables pour tous les types de diffusion spatiale, y compris la propagation des maladies. En ce sens, il fait figure de tenant de la pure analyse spatiale. Dans la postface à *Innovation diffusion as a spatial process*, Pred établit à cet égard une connexion entre Hägerstrand et son prédécesseur suédois Sten de Geer. Pour ce dernier, la géographie devait s'occuper en priorité de « la répartition, de l'expansion, de la distribution des objets, c'est-à-dire de certaines de leurs qualités abstraites, mais pas des objets eux-mêmes dont l'étude relèverait d'autres sciences » (Pred, 1967). L'influence durable qu'exerce *Innovation diffusion as a spatial process* sur la géographie pourrait expliquer qu'au terme de sa revue de littérature sur l'innovation en géographie, Pierre Gondard, alors à l'ORSTOM (Office de la Recherche Scientifique et Technique Outre-Mer), organisme aujourd'hui remplacé par l'IRD (Institut de Recherche pour le Développement), affirme : « l'innovation comme telle n'est sans doute pas un objet de la géographie » (Gondard, *op. cit.*).

Malgré la volonté d'Hägerstrand d'encourager les géographes à étudier la diffusion spatiale d'un large spectre d'innovations – un spectre qui ne se réduise pas au champ des

techniques agricoles – la majorité des travaux relevés par Gondard en 1991 concerne la diffusion des modernisations agricoles, ce qui indique bien, malgré tout, que la discipline a ses objets (*ibid.*). Si le constat de Gondard est peut être biaisé par son appartenance institutionnelle, il est conforté par celui de Michel Chevalier. Organisateur d'un colloque à l'Université de Paris-Sorbonne portant sur « la Géographie de la créativité et de l'innovation » en 1987, ce dernier déclare : « Permettra-t-on au présentateur de terminer sur un sentiment personnel ? Les géographes étudiant les processus de créativité paraissent plus à l'aise dans les secteurs traditionnels de la discipline : le monde agricole et pastoral, le monde semi-rural des bourgs et petites villes. En revanche, la grande ville, qui demeure malgré tout le domaine essentiel de la créativité, met en jeu des facteurs si complexes que, pour leur faire face, la collaboration de l'ensemble des sciences humaines paraît indispensable. » (Chevalier, 1989, p. 8). À notre avis, en dehors de la classique opposition entre géographie rurale et urbaine, il y a un malaise jusqu'aux années 1980 chez les géographes à traiter d'innovations qui ne soient pas techniques et, plus précisément, qui n'aient pas un impact direct sur le paysage ou le milieu naturel. Sans doute, les théories du sous-développement, qui ont cours jusqu'à la fin de la guerre froide, mais aussi l'intérêt pour la transition progressive et généralisée vers une agriculture intensive ont entretenu cette tendance. Mais, nombreux sont les géographes qui admettent l'existence d'un attachement de la discipline à la matérialité. Nous pouvons retenir au moins trois exemples qui en témoignent. Premièrement, dans son propos en faveur d'une géographie des télécommunications, Henry Bakis suggère : « L'une des raisons principales de ce manque d'intérêt [de la géographie, pour les télécommunications,] peut être cherchée dans l'aspect immatériel et caché du phénomène. Homme de terrain, le géographe aime à prendre un contact quasi charnel avec son objet d'étude. Alors que les infrastructures de transports sont visibles, bien ancrées dans le paysage, tangibles même, il n'en va pas de même pour les infrastructures de télécommunications, qui, croit-on dans une première approche, ne s'imposent pas à l'observateur lorsqu'elles sont visibles, et sont même souvent en dehors du champ d'observation (câbles enterrés). » (Bakis, 1980). Deuxièmement, dans l'entrée « production » du dictionnaire *Les mots de la géographie*, Roger Brunet indique que : « la géographie met beaucoup en avant les productions, entendues comme bien matériels » (R. Brunet, 1992, p. 402). Enfin, cette affinité des géographes pour le terrain et l'observation directe des phénomènes est étudiée en détail dans sa thèse par Yann Calbérac, intitulée *Terrains de géographes, géographes de terrain. Communauté et imaginaire disciplinaires au miroir des pratiques de terrain des géographes français du XXe siècle* (Calbérac, 2010).

C'est probablement pour dépasser ce biais dans l'appréhension des innovations, que Gunnar Törnqvist, issu de l'École de Lund et partisan d'une géographie culturelle, introduit l'usage du terme « créativité » en géographie dans les années 1980 (Törnqvist,

1983, 1989, 2011). « Créativité » permet de diriger l'attention sur les innovations scientifiques et artistiques ; en témoigne le fait qu'un spécialiste des études sur la science dans les pays en développement mobilise ce terme en 1979 pour argumenter en faveur de la rupture avec l'approche diffusionniste du développement (Jurdant, 1979). Enfin, cette notion est associée plus facilement au fait urbain que ne l'est alors « innovation ». La théorie très critiquée de la « classe créative » développée par Richard Florida dans les années 2000, tout comme les travaux d'Allen Scott, qui viennent tempérer les propos controversés du premier, relèvent de cet état d'esprit (Florida, 2005a ; Scott, 2006). Toutefois, au moment de fonder une géographie de l'art, Boris Grésillon revient sur ce concept pour en montrer les limites et regretter sa confusion avec celui de « création » qui se prête mieux à la dimension productrice de l'activité artistique (Grésillon, 2014, p. 124-130). Préférer la « création » à la « créativité » ne l'empêche pas de défendre l'existence d'un lien étroit entre l'art et la ville, car la ville peut très bien être considérée comme un lieu de production. Elle est à l'honneur dans le sous-titre de son ouvrage tout juste publié : *Géographie de l'art : ville et création artistique (ibid.)*. Or, un peu comme la notion de « création » a été négligée au profit de « créativité », les spécialistes de l'urbain ont eu tendance, en même temps qu'ils abandonnaient l'approche diffusionniste, à se préoccuper davantage de « hautes technologies » que de « sciences et techniques ». Là encore, puisque les deux phénomènes sont liés, le fait de privilégier le premier témoigne de l'adhésion à un nouveau cadre de pensée pour étudier l'innovation, qui nous vient en grande partie de l'économie. Cette influence explique qu'au lieu d'être technique et rurale, l'innovation soit progressivement devenue technologique et urbaine en géographie. Ce mouvement explique qu'il soit devenu nécessaire de considérer simultanément la recherche et l'innovation, et au-delà de mettre en avant le rôle de la recherche pour le développement des territoires.

2.2. ...à la géographie de l'innovation

L'intérêt pour le lien entre « l'innovation » et « la ville » s'est renforcé dans la géographie des années 1980 sous l'influence des instances d'aménagement du territoire et de la science régionale en pleine expansion. La RERU (Revue d'Économie Régionale et Urbaine) est créée en 1978 par l'Association de Science Régionale de Langue Française, ouverte aux disciplines de l'économie, la géographie, l'aménagement, la gestion, la sociologie, et les sciences politiques (Bailly, Derycke, & Torre, 2012). De plus, au début des années 1980, le GREMI (Groupe de Recherche Européen sur les Milieux Innovateurs) se constitue sous la houlette de Philippe Aydalot (Tabariés, 2005). Ce groupe est surtout formé d'économistes qui s'inscrivent dans la lignée de François Perroux et rediscutent la

notion d'« économie externe » forgée en 1890 par Alfred Marshall, et de « pôle de croissance » développée par Perroux (Aydalot, 1965). En réaction à la crise industrielle, les spécialistes en science régionale et urbaine adhèrent à une définition qui présente l'innovation comme une source de richesse ou une fin « en soi ». Ce qui importe dans cette perspective n'est plus tellement l'adoption d'innovations, mais la capacité des hommes à innover. Dans les modèles de croissance économique, cela se traduit par une intégration du changement technologique comme variable endogène (Romer, 1986). À l'inverse, dans les modèles dominants des années 1950-60 comme celui de Solow (1957), le changement venait de l'extérieur (facteur exogène) (Romer, 1990). En fait, comme beaucoup de progrès techniques et médicaux ont été réalisés dans l'immédiat après-guerre grâce à la sphère publique (et plus précisément aux organismes de recherche publique), il était de rigueur de considérer jusqu'aux années 1980 que les acteurs économiques pouvaient en profiter librement.

À partir des années 1980, on considère que l'État providence n'est pas le seul acteur à pouvoir faire progresser la société. Les nouveaux modèles de croissance endogène rencontrent de l'audience parmi les géographes, en particulier les spécialistes de géographie économique. Dans sa récente thèse, Antoine Grandclément considère l'évolution qui en résulte – qu'il décrit avec une grande précision – comme le passage en géographie industrielle « d'une géographie de la production à une géographie de l'innovation » (Grandclément, 2012, p. vi). Le constat de départ est que les capacités à innover ne sont pas réparties équitablement dans l'espace géographique et que parmi l'ensemble des lieux, seule une minorité semble disposer de moyens effectifs pour innover. Or, parce que les activités économiques et les hommes s'y concentrent, les grandes villes, celles qui au XX^e siècle ont gagné un nombre considérable d'habitants en un court laps de temps, sont envisagées comme des entités spatiales propices à l'apparition d'innovations. Cette idée d'une synergie propre à la concentration des hommes et des activités en ville est ancienne en géographie (Robic, 2003) mais elle devient centrale au moment où l'urbain s'affirme en tant que champ d'investigation interdisciplinaire et lieu de la reprise économique.

D'une manière générale et par-delà les réflexions internes au domaine hybride que constitue la science régionale et urbaine, l'intérêt pour le lien entre « l'innovation » et les grandes villes qui, puisqu'elles regroupent beaucoup d'activités tertiaires voire quaternaires (les fonctions de commandement et les services rares) sont qualifiées de « métropoles », n'a fait que progresser depuis le début des années 1980 en géographie urbaine, économique et industrielle (Benko, 2007 ; Daviet, 2005). Cette attention se développe progressivement et parallèlement à une préoccupation pour les « districts » ou « clusters » qui sont des zones caractérisées par une synergie particulière entre les divers acteurs économiques qui y cohabitent, améliorant leur capacité à innover, et donc à générer des béné-

fices. C'est à ce moment que l'utilisation de notions comme « territoire » et « milieu » entre en concurrence avec celle du terme « espace » (pour faire simple, l'« espace » est un terme trop neutre et creux à l'heure du « tournant culturel »). Il faut préciser que sur le terrain de l'action publique, cette évolution coïncide avec la décentralisation politique qui cherche à responsabiliser les entités spatiales infra-étatiques. Ce processus donne de la consistance à l'idée de « territoires », parce que contrairement à l'« espace », le « territoire » dispose d'une capacité d'action et d'une autonomie. Petit à petit, le principe du « développement local » vient bousculer une vision plus classique de l'aménagement du territoire national, portée en France par la DATAR (Délégation à l'Aménagement du Territoire et à l'Action Régionale) – actuelle CGET – et axée sur l'égalité des territoires. Au lieu d'agir uniquement par des mécanismes de péréquation et de redistribution à l'échelle nationale, les politiques publiques se fixent pour nouvel objectif d'inciter et de favoriser la dynamique et l'autonomie économique locale, donc les capacités d'innovation des territoires infranationaux et plus particulièrement, des régions et des villes.

Dans ce contexte, certains lieux dotés d'équipements scientifiques sont pris comme modèles de réussite en matière d'innovation et de croissance endogène. Les géographes sont incités à se pencher sur ces espaces jugés prometteurs. Parmi les plus connus, citons la Silicon Valley dans la Baie de San Francisco et la Route 128 qui longe Boston (Brocard, 1978). La capacité de ces lieux à supporter le désengagement financier de l'État et les reconversions industrielles provoque l'admiration et leur dynamisme scientifique semble y contribuer (Soppelsa, 1976). Politiquement, on arrive au stade – déjà évoqué dans la section précédente (Section 1) – où les États se soucient de rendre plus efficace la répartition de leurs dépenses en matière de recherche. La mesure des débouchés et des résultats de la recherche en lien avec le potentiel d'innovation des territoires intéresse de plus en plus les responsables politiques. Des opérations d'aménagement sont lancées pour édifier des « technopôles » ou « parcs scientifiques, technologiques et d'affaire », c'est-à-dire des lieux qui, comme la Silicon Valley, concentrent des équipements publics de recherche ainsi qu'un tissu d'entreprises innovantes (George, 1987 ; Mérenne-Schoumaker, 1991). Alors que les géographes s'intéressent à ces nouveaux ensembles, alternativement appelés « clusters », « districts technologiques » ou « technopôles », qui symbolisent l'entrée dans une nouvelle ère, une ère « postfordiste » propice au développement des activités de « haute technologie », la thématique de la grande ville n'est jamais très loin. Par exemple, en 1987, Pierre George établit un lien entre le technopôle et la ville du futur qualifiée de « technopole » : « Dans la mesure où les nouvelles formes d'industrie à haute productivité, avec leur accompagnement de services scientifiques, s'implantent dans des lieux nouveaux, en annexe à des villes dont les fonctions antérieures leur étaient étrangères, on voit surgir un nouveau type de relation entre l'industrie et le site au sens

global du terme : le *technopôle* donne naissance à une nouvelle forme de ville occupée par un nouveau type de population active, la *technopole*. » (George, *ibid.*). Les études de cas se multiplient qui témoignent de cette nouvelle approche de l'urbain (Jalabert & Grégoris, 1987 ; Benko, 1990). L'interprétation dominante est qu'à la grande entreprise fordiste et ses « économies d'échelle » succéderait un système friand en innovations s'appuyant sur des échanges entre firmes, mais aussi entre science et industrie, et plutôt fondé sur des « économies d'agglomération ». En 1984, l'analyse de Michael J. Piore et Charles Sabel sur l'avènement d'un régime de production laissant davantage de place à la « flexibilité » tend à s'imposer (Piore & Sabel, 1984). George Benko résume ainsi la nouvelle orthodoxie : « du plus petit district italien aux mégapoles mondiales, le nouveau paradigme technologique de la "spécialisation flexible" impulserait non seulement le retour des usines et des bureaux vers les zones urbaines, mais encore la reprise de la croissance quantitative des métropoles : forme spatiale enfin trouvée de sortie de la crise du fordisme. La future hiérarchie des villes et régions urbaines mondiales résulterait de la stratégie interne de ces districts (ou amas de districts) : Que les meilleurs gagnent ! » (Benko & Lipietz, 1992 ; Benko, 2007). Au cours des années 1990, la statistique publique entérine cette approche renouvelée de la société, ce dont témoigne la création à l'INSEE (Institut National de la Statistique et des Études Économiques) d'une nouvelle catégorie d'activités qui recoupe le secteur tertiaire pour regrouper les fonctions « stratégiques », bientôt qualifiées de « fonctions métropolitaines supérieures » (Julien & Pumain, 1996 ; Julien, 2002). Sans hésitation, la recherche est promue au rang d'activité métropolitaine, ce qui contribue en retour à orienter son traitement géographique (Saint-Julien, 1999).

En 1992, dans un ouvrage sur la politique de développement des pôles technologiques, Gabriel Wackermann formule le bilan suivant pour la France où l'opposition entre Paris et la Province est un enjeu structurant : « de 1954 à 1975, l'objectif de la politique d'aménagement du territoire consistait à mieux répartir la population et les activités démesurément concentrées dans la région parisienne. De nos jours encore, l'Ile-de-France participe puissamment au déséquilibre national mais Paris n'est plus perçu comme un obstacle au déploiement des régions françaises : la capitale sert d'atout, de mégapole utile dans le cadre de la compétition internationale. Dans cette perspective, l'Ile-de-France s'est dotée d'une fraction significative du potentiel d'industries de pointe et de recherche disponible en France. (...) en dépit de ce poids parisien, la province a accru sa propre part dans les emplois profitant de la politique de décentralisation d'entreprises largement encouragée. À la conception égalitariste de l'aménagement fondée sur un souci d'équité et de cohésion sociale s'est substituée une conception darwinienne. La première devait aider les régions les moins équipées à rattraper leur retard par rapport aux régions privilégiées en provoquant des transferts interrégionaux d'attributions budgétaires. La seconde con-

siste à hisser un nombre aussi important que possible de régions au rang des espaces suffisamment structurés pour conduire la compétition internationale. » (Wackermann, 1992). Dans ce cadre théorique, la recherche scientifique, au même titre que d'autres activités favorisant l'innovation, accompagne la croissance économique dont le siège se trouve désormais en ville et plus précisément, dans les espaces métropolitains. Toutefois, le contenu de la science, son inscription dans des cadres institutionnels structurants de même que son lien à l'enseignement restent peu envisagés. Quelques géographes s'intéressent aux restructurations de la carte universitaire (voir Chapitre 3.1.2) mais cette approche apparaît comme déconnectée de l'approche par l'industrie et l'innovation. Seule Madeleine Brocard, qui a pris pour objet « la recherche scientifique » est à la jonction entre ces deux courants, comme nous le verrons dans le chapitre suivant (Brocard *et al.*, 1996).

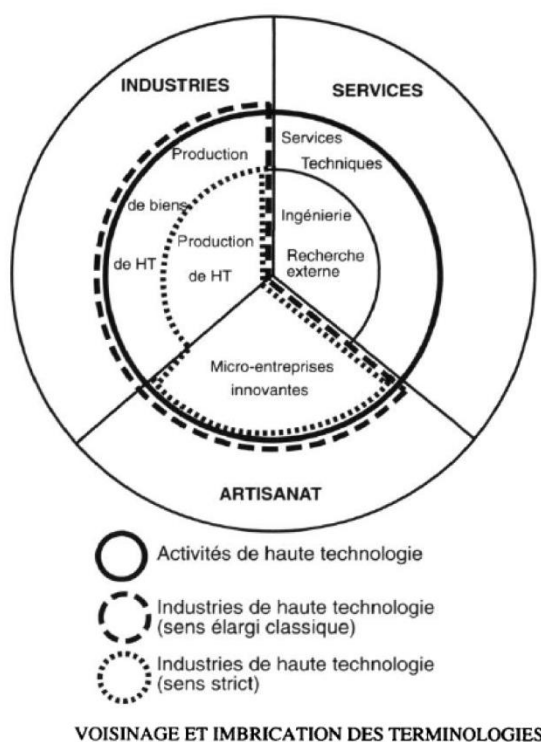


Figure 5 — *Voisinage et imbrication des terminologies. Schéma extrait de l'article « La définition des industries de haute technologie » (Fache, 1999)*

Du côté de l'approche par la géographie de l'industrie, des lieux peuplés de scientifiques ainsi que les régions-cibles des investissements en recherche et développement font l'objet d'études de cas approfondies, surtout pour saisir comment les activités scientifiques peuvent favoriser le développement économique local, mais l'insertion effective de ces lieux dans des réseaux scientifiques nationaux et internationaux, de même que les flux de connaissance entre territoires, ne sont pas étudiés par cette géographie avant le milieu

des années 2000. Aussi, dans les travaux portant sur les industries de haute technologie, c'est moins la géographie de la recherche de pointe qui est interrogée que la localisation des lieux de production de biens nouveaux dont la conception résulte des efforts de recherche (Fache, 1999). Le flou terminologique autour de la notion de « haute technologie » légitime toutefois des approches transversales et une prise en compte de la recherche, au moins appliquée, dans l'analyse géographique (Figure 5). Et, en raison des rapports de plus en plus privilégiés entre recherche technologique et industrie, Solange Montagné-Villette en vient à envisager l'idée d'une dilution de la frontière entre le secondaire et le tertiaire (Montagné-Villette, 2001).

Parallèlement, mais sans être vraiment déconnectée de cette géographie, la science régionale et urbaine contribue à enrichir l'approche spatiale de l'innovation. Progressivement, à la focalisation sur la concentration des activités et les relations qui se tissent au sein des « milieux innovateurs » (l'expression est de Philippe Aydalot), des systèmes nationaux (Patel & Pavitt, 1994) et régionaux d'innovation (Asheim & Coenen, 2005), l'attention s'oriente vers l'inscription des territoires de l'innovation dans des systèmes d'échanges, pas uniquement marchands, à plusieurs niveaux scalaires. Cette évolution est contemporaine des réflexions en science sociales sur les rapports entre « réseaux » et « territoires » que le sociologue Manuel Castells a encouragées (Castells, 1996). Pour simplifier et généraliser, il semblerait qu'alors que dans les années 1980, on pouvait opposer d'un côté, les études démontrant un souci croissant pour la dimension locale et endogène des phénomènes et, de l'autre, les travaux structuralistes ayant l'ambition de construire une théorie générale des dynamiques mondiales ; il ait fallu attendre les années 1990 pour que le besoin de concilier approche locale et globale des phénomènes s'impose (Amin & Thrift, 1992). En effet, à l'aube de la décennie, le débat fait rage entre plusieurs approches du développement des territoires. Aux études de cas allant des « districts industriels italiens » aux « clusters californiens », des chercheurs plus structuralistes opposent une vision de l'organisation des territoires dans laquelle l'État et les grandes entreprises jouent toujours un rôle décisif et dans laquelle le moteur du développement local serait social avant d'être économique. Le recueil publié par Georges Benko et Alain Lipietz en 1992 intitulé *Les Régions qui gagnent, districts et réseaux : les nouveaux paradigmes de la géographie économique* fournit un intéressant éventail des positions en présence dans ce débat (Benko & Lipietz, 1992). Dans cette dynamique, l'approche sociologique gagne de nombreux adeptes. Il y a pendant cette période, un souci de plus en plus marqué pour les mécanismes de coordination et les logiques d'organisation institutionnelles (Chabault, 2006). Comme on va le voir dans le chapitre suivant, les apports de la sociologie économique, et notamment le thème de l'« encastrement » des relations sociales au marché économique, sont progressivement assimilés par l'économie régionale (Ferru, 2009, p. 71-81). C'est

grâce à cette influence grandissante des théories sociologiques, sur laquelle nous reviendrons à travers l'exemple de la ligne de recherche toulousaine impulsée par Michel Grossetti, qu'il va être possible de concilier plutôt que d'opposer approches endogènes et exogènes de l'espace ; diffusionnisme et environnementalisme ; rôle de l'État et des grosses entreprises dans l'aménagement des territoires d'un côté, et capacité des territoires à s'auto-organiser grâce à des stratégies de développement local, de l'autre.

C'est aussi dans cette mouvance qu'en géographie, l'attention pour des cas particuliers de technopoles est progressivement remplacée par un souci renouvelé pour les réseaux de villes et, plus particulièrement, pour les réseaux de grandes villes comme en témoigne le terme d'« Archipel Mégalopolitain Mondial » (AMM) qu'Olivier Dollfus définit comme « un ensemble de villes qui contribuent à la direction du monde, (...) un symbole de la globalisation liée à la concentration des activités d'innovation et de commandement » (Dollfus, 1997, p. 33). Que ce soit par le biais d'une entrée monographique ou réticulaire, le sens associé par les géographes à la notion d'innovation s'est donc indéniablement transformé à partir des années 1980, si bien que lorsqu'Yves Guermond se charge en 2001 de revenir sur le traitement géographique de l'innovation, il part de Mark Jefferson mais ignore la période pendant laquelle les innovations étaient essentiellement perçues comme techniques et issues du milieu rural (Guermond, 2001). Hormis le travail de Jefferson, il ne cite que des travaux réalisés après 1980, mais les références sont très hétérogènes, ce qui souligne le caractère hétéroclite de l'objet « innovation » en géographie. Le contenu précis des activités innovantes ne semble pas avoir d'importance puisque Guermond met sur le même plan des travaux de géographie sur les dépôts de brevets, la chimie des polymères, les services aux entreprises, les activités multimédias et les contrats de recherche CNRS-entreprise. Dans l'ensemble, il s'agit d'étudier l'inégale répartition des activités innovantes et, en particulier, leur concentration géographique. Guermond souligne qu'entre autres facteurs, l'accessibilité et la connectivité des territoires augmentent leur propension à accueillir des activités innovantes. Cette approche de l'objet « innovation » est économique en ce qu'elle associe l'innovation aux capacités d'enrichissement des territoires, ce qui confirme l'influence de l'économie des territoires sur la géographie contemporaine. À la fin de cet article, la présence de chercheurs et la délocalisation des laboratoires de recherche est évoquée. Guermond considère que ces variables sont fondamentales pour évaluer le potentiel d'innovation d'un territoire mais il conclut sur la difficulté d'évaluer les politiques territoriales de recherche parce que les transferts de connaissance nécessaires à l'innovation et la qualité de cette dernière sont difficilement mesurables.

CHAPITRE 3. VERS UNE GÉOGRAPHIE DES SCIENCES

Le constat suggérant qu'il est difficile de mesurer les mouvements de connaissances et par là, d'évaluer l'efficacité des politiques territoriales de recherche garde sous silence le fait qu'entre la fin des années 1970 et les années 2000, le cadre méthodologique permettant de traiter l'innovation s'est étoffé et qu'un domaine de recherche transversal, appelé « sciences de l'innovation », s'est constitué. L'une de ses principales préoccupations est d'étudier la dimension spatiale de l'innovation, et plus précisément, de saisir les « externalités technologiques » dont profitent les acteurs économiques agglomérés à proximité d'unités de recherche et au-delà, la circulation spatiale des connaissances. Parallèlement, un domaine d'étude spécialisé s'est développé autour de la géographie de l'enseignement supérieur qui a donné de plus en plus de place aux questions relatives à l'évaluation et l'efficacité des unités de recherche et de formation. À partir des années 2000, le dialogue difficile entre ces deux domaines, qui ont pourtant en commun le thème de la recherche et de la mesure de l'activité scientifique localisée, soulève des problèmes relatifs à l'organisation des systèmes urbains. En effet, les activités innovantes et les activités de formation ne sont pas supposées obéir aux mêmes logiques spatiales. Cette problématique offre, c'est l'objet du présent chapitre, une opportunité intéressante pour que s'opère une intégration plus réfléchie et entière de l'objet « science » en géographie. Toutefois, en dehors du cadre des *science studies*, les questions relevant de la géographie de la recherche restent encore souvent, on va le voir, envisagées en lien avec le potentiel d'innovation des territoires ou bien, indirectement, par le biais de l'offre d'enseignement supérieur sous l'angle de l'aménagement de la carte universitaire. Or, dans le cadre d'une géographie des sciences, il serait pertinent de développer une analyse de la répartition des activités de recherche et de leur fonctionnement qui ne se réduise ni à une géographie de l'innovation, ni à une géographie de l'enseignement supérieur mais qui tienne également compte des logiques proprement scientifiques qui caractérisent cette activité. Les évolutions intervenues au cours des années 1990-2000 dans le cadre de la géographie de l'innovation, de la géographie de l'enseignement supérieur et de la géographie des systèmes urbains sont néanmoins un pas en avant pour une géographie des sciences.

1. La géographie de la recherche : entre logiques d'innovation et de formation

« Or, il nous est apparu peu à peu que ce que l'on appelait développement ne recouvrait en réalité que des processus de croissance et que le contenu même du terme recherche était souvent mal perçu ; surtout depuis que l'on parle beaucoup d'innovation. » (Brocard, 1978, p. 1-2)

1.1. Au-delà d'une géographie de l'innovation...

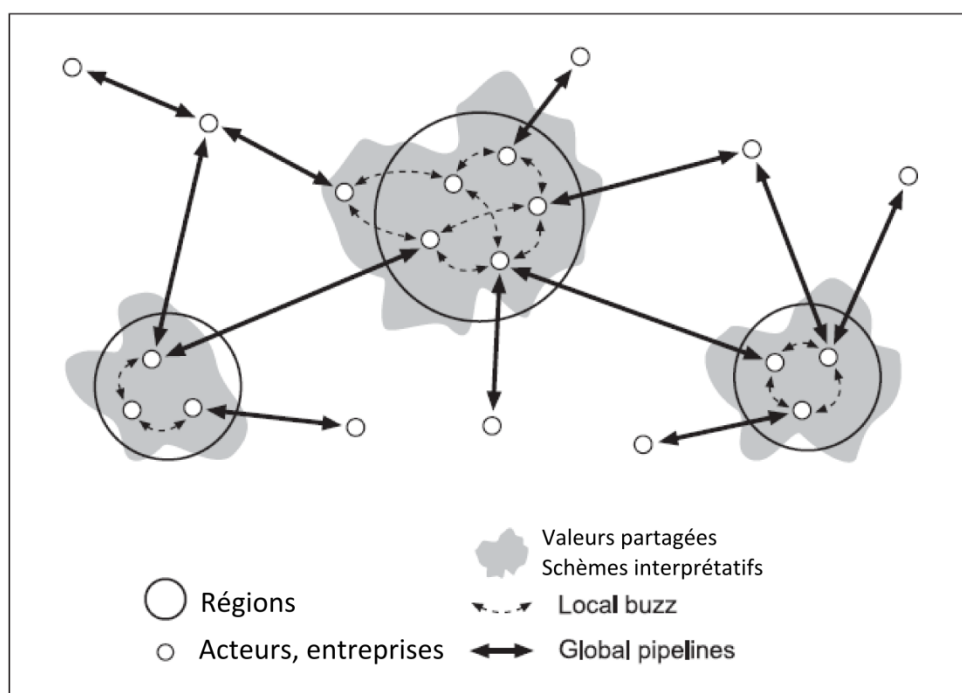
Les recherches qui traitent des bienfaits de la concentration spatiale des activités humaines ont gagné en visibilité et en attractivité suite au succès de la théorie économique de Paul Krugman au début des années 1990 qui lui a valu un prix de la Banque centrale de Suède en sciences économiques en mémoire d'Alfred Nobel (en 2008) dont « géographes et économistes se félicitent et se disputent plus ou moins violemment la paternité et les filiations » (Lacour, 2008, p. 7). Même si pour certains, Krugman n'a fait que réinventer l'eau chaude en géographie, son travail a participé au développement d'une « Nouvelle Économie Géographique » (Klein & Guillaume, 2014), sans laisser indifférent les géographes (R. Martin, 1999 ; « Nouvelle économie géographique et géographie : Quel dialogue ? », 2007). Pour les spécialistes de l'innovation, il y a un gros problème dans l'approche de Krugman puisqu'il souligne l'importance des « effets d'agglomération » pour la croissance économique, mais néglige le rôle des transferts de connaissance qu'il présente comme des impondérables (Autant-Bernard, Billand, & Massard, 2010). Or, il existe un certain nombre d'économistes qui défendent, depuis la fin des années 1970, qu'en plus d'être mesurables, les flux de connaissances méritent d'être analysés de près, notamment pour déterminer le potentiel d'innovation qui dérive de la concentration spatiale des facteurs de production. Ce sont ces chercheurs, souvent issus de l'économie industrielle dans le cas français, qui sont à l'origine d'une nouvelle branche de l'économie qualifiée de « géographie de l'innovation » (Autant-Bernard & Massard, 2001). À l'origine de cette branche de recherche qui s'attache à expliquer la localisation des innovations, on trouve des économistes américains qui s'inscrivent dans l'héritage de Zvi Griliches. À la suite de celui-ci, ils travaillent à l'aide d'une fonction de production de connaissances (à ne pas confondre avec une fonction de production économique classique) qui intègre le facteur technologique (Griliches, 1979).

Pour modéliser les transferts de connaissances, également appelés externalités ou *spillovers* (littéralement « débordement ») technologiques, les économistes américains ont la bonne idée de mobiliser des données scientométriques. À partir des années 1980 en particulier, un post-doctorant de Griliches à Harvard, Adam Jaffe, considère des données du bureau des brevets américains et du NSF pour mesurer l'effet local de la recherche académique sur le potentiel d'innovation des entreprises. Il questionne notamment le lien existant, au niveau régional, entre un indice de « coïncidence géographique » (ou co-présence) des universités avec les industries (*input*) et le nombre de dépôts de brevet (*output*) (Jaffe, 1989). Dans les années 1990, Zoltan Acs, Maryann Feldman et David Audretsch poursuivent dans cette voie et montrent que la recherche académique n'est pas la seule responsable d'externalités : les efforts de recherche et développement des établissements privés génèrent, eux aussi, des débordements ou externalités de connaissance dont peuvent profiter leurs voisins (Acs, Audretsch, & Feldman, 1992 ; Audretsch & Vivarelli, 1994 ; Anselin, Varga, & Acs, 1997). Sur ces acquis et ceux d'Edwin Mansfield, Feldman en vient à promouvoir le développement d'une spécialité de recherche tournée vers l'étude des déterminants locaux de l'innovation. Bien qu'elle mobilise peu de géographes, Feldman baptise cette spécialité : « geography of innovation » (Feldman, 1994). En France, les économistes Jacques Mairesse, Nadine Massard et Corinne Autant-Bernard en sont les principaux importateurs (Autant-Bernard, Mairesse, & Massard, 2007). Alors que dans les années 1970 et 1980, le lien entre l'investissement dans la recherche et le dynamisme économique local était postulé pour légitimer l'aménagement de parcs scientifiques ou technopôles, les travaux qui se réclament de cette géographie de l'innovation traduisent une volonté de mieux capturer les mécanismes qui se cachent derrière ce lien (Varga, 1997 ; Autant-Bernard, 2001).

À ce titre, un dialogue s'instaure entre les spécialistes français de l'innovation et les membres du groupe de recherche « dynamiques de proximité » qui se constitue pendant les années 1990. Ce groupe issu, lui aussi, de l'économie industrielle entend nuancer le rôle de la proximité géographique dans la circulation des connaissances. L'argument majeur est que la proximité géographique s'accompagne de proximités organisationnelles ou « organisées » qui sont, quant à elles, déterminantes pour la circulation du savoir nécessaire à l'innovation (Rallet & Torre, 1999 ; Zimmermann, 2008 ; Bouba-Olga & Grossetti, 2008 ; Torre, 2008 ; Carrincazeaux, Lung, & Vicente, 2008 ; Talbot, 2009). Dans le même esprit que ce groupe français, un groupe de chercheurs néerlandais étudie la géographie de l'innovation à l'aide d'une grille d'analyse développée par Ron Boschma, qui distingue cinq types de proximité (géographique, organisationnelle, institutionnelle, cognitive et sociale) associés à des mécanismes spécifiques d'appariement entre les acteurs de l'innovation (Boschma, 2005 ; Balland, 2011 ; Ter Wal, 2014). Cette évolution est révélé-

latrice de la place plus grande accordée à la sociologie par les économistes soucieux de la dimension spatiale des phénomènes. Les approches par la « proximité » se justifient d'autant plus que les relations à distance – les échanges et flux d'hommes, d'information et de connaissance – entre les territoires se sont révélées essentielles dans plusieurs études empiriques sur l'innovation (Jaffe, Trajtenberg, & Henderson, 1993 ; Almeida & Kogut, 1999 ; Autant-Bernard, 2001 ; Saxenian, 2006).

En effet, l'image du territoire comme espace fermé doté d'une dynamique endogène entretenue par les vertus de la seule proximité géographique entre les acteurs ne résiste pas aux observations empiriques (R. Martin & Sunley, 2003). À la suite d'échanges entre économie, géographie et sociologie, un nouveau modèle explicatif l'emporte (Figure 6). Il est décrit pour la première fois dans l'article publié en 2004 par les géographes scandinaves Harald Bathelt, Anders Malmberg et Peter Maskell : « *Clusters and knowledge : local buzz, global pipelines and the process of knowledge creation* » (Bathelt, Malmberg, & Maskell, 2004). Pour éviter le *lock in* ou l'enfermement sur le milieu local qui est potentiellement sclérosant (Broekel & Boschma, 2012), et faire en sorte que les lieux de l'innovation soient capables de « résilience », les spécialistes admettent qu'il faut privilégier les échanges de connaissances entre territoires (Bathelt, 2007 ; Suire & Vicente, 2012).



Les dynamiques réflexives des "local buzz" et "global pipelines"

Figure 6 – Les dynamiques réflexives des « local buzz » et « global pipeline ». Schéma extrait et traduit de l'article « *Buzz-and-Pipeline Dynamics: Towards a Knowledge-Based Multiplier Model of Clusters* » (Bathelt, 2007)

Ce modèle visant à expliquer les phénomènes d'innovation rappelle celui des réseaux de transport aérien de type « *hubs and spoke* » (de Roo, 1993 citée par Mérenne-Schoumaker, 1996 ; O'Kelly, 2008) qui, comme modèle d'organisation de l'espace, tend à remplacer la théorie des lieux centraux à partir des années 1990. Il s'agit aussi du principe d'organisation théorisé par le sociologue Manuel Castells pour qui, grâce aux possibilités des nouvelles technologies d'information et de communication, « l'espace des flux » devient l'espace de concentration du pouvoir, notamment économique, au détriment de « l'espace des lieux » (Castells, 1996 ; Stock, 2011). Dans un contexte où la circulation des biens immatériels est facilitée par l'accès de plus en plus massif aux nouveaux moyens d'information et de télécommunications (NTIC), le succès de ces modèles réticulaires se traduit pendant les années 2000 par l'adoption des méthodes de l'analyse des réseaux sociaux en géographie économique (Ter Wal & Boschma, 2009).

À partir des années 2000, l'intérêt croissant pour les flux de connaissances et la valorisation du rôle de la science dans l'intégration territoriale permettent enfin aux réseaux de collaborations scientifiques d'entrer dans le champ d'analyse des spécialistes de l'innovation. On repère alors une volonté épistémologique qui s'accompagne d'une volonté politique de dépasser la « logique *inputs/outputs* ». Nadine Massard, qui fait partie des membres structurants de la géographie de l'innovation en France, s'en explique dans un livre écrit avec Sylvie Chalaye : « Les tableaux de bord localisés existants restent souvent limités à la production d'indicateurs *inputs/outputs* comparant strictement ou associant moyens mis dans la recherche d'un côté et résultats en terme d'innovations repérables de l'autre côté. Il manque très généralement une grille essayant de mettre en avant les principales caractéristiques structurelles et mécanismes organisationnels, collectifs, interactifs de l'innovation susceptibles d'expliquer pourquoi, à moyens mis dans la recherche équivalents, certaines régions sont plus innovantes que d'autres. » (Chalaye & Massard, 2012, p. 11). Poussés par les impératifs européens et internationaux, les travaux se multiplient qui tâchent d'analyser les mécanismes de l'économie de la connaissance et d'évaluer le succès des politiques publiques touchant à la recherche et au développement. En plus des études classiques portant sur la distribution des dépenses régionales de recherche et développement ou sur les dépôts et citations de brevets (Sterlacchini, 2008 ; Paci & Usai, 2008 ; Breschi & Lissoni, 2009), apparaissent des études spatialisées portant sur les Programmes Cadre de Recherche et Développement (PCRD) européens ou plus récemment encore, au niveau français, sur les pôles de compétitivité (Hoekman, Frenken, & Oort, 2009 ; Frachisse, 2011 ; Levy, Hussler, & Triboulet, 2012). Dans ce domaine, la science qui débouche sur des publications académiques fait l'objet d'une attention plus marginale parce que son rapport à l'activité économique est moins direct. Mais, les exceptions tendent à se multiplier ces dernières années et l'on peut penser, par exemple, aux

thèses de Roderik Ponds (Ponds, 2008) et Jarno Hoekman (Hoekman, 2012) ou, en France, à celles de Marie Ferru (Ferru, 2009), Mickaël Benaim (Benaim, 2013) et Bastien Bernela. Si l'innovation demeure le souci principal de cette spécialité, la science n'est pas complètement absente puisqu'une session spéciale « géographie de la science » s'est tenue à chaque édition de la conférence biannuelle européenne de géographie de l'innovation depuis la première en 2010 (2010 à Toulouse, 2012 à Saint-Étienne, 2014 à Utrecht, puis à venir 2016 à Toulouse). Cette session a d'ailleurs été l'occasion de communiquer les progrès réalisés par notre équipe en matière de spatialisation des données bibliométriques.

En dépit de l'intégration croissante de l'objet « science » dans cette géographie de l'innovation, les limites de l'approche par l'« innovation », déjà pointées par Michel Grossetti *et al.*, dans le rapport PIR-villes paru en 1996 sur « la localisation des institutions d'enseignement supérieur et de recherche en France » conservent une certaine actualité : le centrage sur la dimension spatiale de l'innovation conduit à ignorer plusieurs dimensions fondamentales de l'activité scientifique (Grossetti *et al.*, 1996). Premièrement, la dimension historique de la spatialisation des établissements de recherche est peu considérée (même si la récente volonté d'intégrer les apports de l'économie évolutionniste dans les analyses témoigne d'une certaine capacité d'adaptation chez les spécialistes de l'innovation, voir (Boschma & Frenken, 2006 ; Balland, Boschma, & Frenken, 2013) ; deuxièmement, la richesse et l'hétérogénéité des activités scientifiques qui dépendent de logiques disciplinaires variées avec des fonctionnements propres – les rythmes et l'organisation du travail différent d'une discipline à une autre, voire d'une spécialité scientifique à une autre – ne sont pas toujours abordés, au prétexte de mettre à jour les lois générales qui régissent les rapports entre science et innovation en privilégiant une approche économétrique ; et troisièmement, les liens institutionnels entre activités de recherche, activité de formation et activités industrielles sont rarement envisagés de concert. Quand le lien entre formation et recherche est considéré par les spécialistes de l'innovation, c'est forcément de façon périphérique tandis que les cas particuliers de relations entre universités et entreprises sur lesquels se fondent la plupart des hypothèses tiennent plus du fantasme que de la réalité : « l'idéal type [implicite qui guide la recherche sur les relations entre les institutions de recherche et les entreprises] peut être décrit comme suit : une équipe de recherche pour ou en collaboration avec une grande entreprise ou une PME de haute technologie du même domaine et qui possède le personnel et les expertises nécessaires à une relation fructueuse ainsi qu'à l'utilisation optimale des résultats obtenus. Or cet idéal type ne permet pas de penser adéquatement les relations entre les institutions de recherche et les entreprises ou, en d'autres termes, les relations entre les différents acteurs de l'innovation. » (Trépanier & Ippersiel, 2003).

Ainsi, dans ce cadre théorique, malgré les efforts qui ont été faits pour une meilleure appréhension empirique des mécanismes de l'innovation, notamment à l'aide de données scientométriques, la science reste à quelques exceptions près, appréhendée comme un moyen de générer des innovations en vue, si possible, d'entretenir la croissance ou la résilience des territoires. Dans cette perspective, les raisons et logiques internes à la pratique scientifique et leur manifestation dans l'espace importent peu. D'ailleurs, l'approche économétrique et l'entrée par l'innovation ne rendent possible qu'une approche partielle des données sur la science. En effet, l'hétérogénéité spatiale tend à être laissée pour compte par le recours à la modélisation. Les spécificités locales sont gommées pour faire entrer les observations empiriques dans des modèles opérants toutes choses égales par ailleurs.

Cartographie des réseaux du pôle PASS à l'échelle régionale

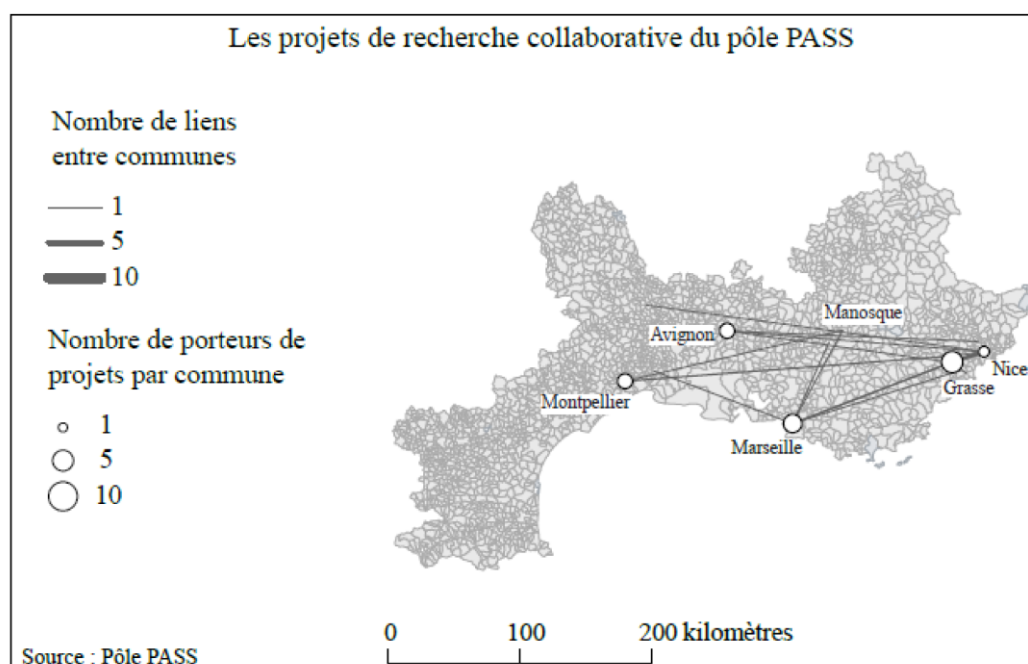


Figure 7 – Cartographie des réseaux du pôle PASS à l'échelle régionale. Schéma extrait de la thèse « Géographie des pôles de compétitivité : réseaux et territoires de l'innovation » (Grandclément, 2012, p. 530)

C'est ainsi qu'en dépit de l'existence de cette géographie de l'innovation, une approche plus géographique des jeux de données spatialisés sur la science et l'innovation demeure pertinente car, loin de vouloir se substituer à ou invalider les travaux que nous venons d'évoquer, elle est en mesure d'apporter un éclairage différent et complémentaire sur la science et l'innovation. Entre autres études récentes qui prennent cette direction en

géographie, la thèse de Grandclément sur les pôles de compétitivité, inscrite dans la géographie de l'industrie, mérite, à nouveau, d'être évoquée parce qu'elle fait la jonction entre approche économique et approche géographique de l'innovation (Grandclément, 2012). Le recours à des méthodes descriptives dans cette thèse comme les représentations graphiques de réseaux et de cartes permettent de faire ressortir l'hétérogénéité spatiale des relations nouées au sein des pôles de compétitivité et les jeux d'échelle, c'est-à-dire l'inscription des réseaux de collaboration scientifiques à plusieurs niveaux géographiques (Figure 7). De même, la thèse de Josselin Tallec qui, pour quatre petites villes françaises, réinscrit les rapports entre territoires et innovation dans le temps long, invite à nuancer le caractère souvent présenté comme novateur du rôle de l'innovation dans la dynamique socio-économique des petites villes (Tallec, 2014). On peut aussi évoquer les thèses antérieures de Bertrand Moro et Marie-Noëlle Comin qui suggèrent, depuis le milieu des années 2000, que la géographie française est entrée dans une phase de valorisation du rôle de la recherche pour l'aménagement du territoire à l'échelle régionale (Moro, 2006) et pour la mise en réseau et la dynamique économique des villes à l'échelle européenne (Comin, 2009).

À la fois pour expertiser les plans en matière d'aménagement de la recherche ou mieux appréhender le dynamisme des sites scientifiques dans une démarche de développement local, les évolutions des années 2000 signalent qu'une géographie des activités scientifiques est requise. Celle-ci pourrait venir enrichir et prolonger les réflexions déjà menées sur l'aménagement de la carte universitaire française, abordées dans la section suivante. L'aménagement actuel de cette carte suggère qu'il y a actuellement trop d'idéologies et peu de savoirs empiriquement fondés sur lesquelles s'appuyer pour aménager ou gouverner la recherche, y compris à cause de l'analogie entre les logiques de production scientifique et les logiques de production économique. Dans les années 2000, le fait que les décideurs français aient pris appui sur le « classement de Shanghai » alors qu'il repose sur des indicateurs fort discutables pour justifier des politiques de recherche ou des stratégies de sites universitaires (Gingras, 2008 ; Théry, 2009 ; Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015) est un révélateur parmi d'autres de la « méconnaissance du fonctionnement réel de l'activité scientifique » qui règne dans les milieux où les décisions sont prises et légitimées (Milard, 2012).

1.2. ...et d'une géographie de l'enseignement supérieur

Quand les activités de recherche ne sont pas abordées à travers leur participation à l'innovation et au développement local, elles sont évoquées, plus ou moins directement, par le biais de leur association à l'enseignement supérieur. En France, la géographie de

L'enseignement supérieur est un domaine qui, comme la géographie de l'innovation, rassemble des chercheurs aux approches variées (d'une géographie sociale à une géographie des systèmes formation-emploi en passant par une géographie des systèmes urbains). Dès lors, même s'il y a finalement peu de géographes impliqués dans ce domaine de recherche, il n'est pas aisé d'en synthétiser les apports. En fait, les réflexions politiques en matière d'aménagement universitaire assorties de demandes d'études de la part de la DATAR (Délégation à l'Aménagement du Territoire et à l'Action Régionale) ou d'institutions comme le PUCA (Plan Urbanisme Construction Architecture, successeur du Plan Urbain) ont obligé, à intervalles plus ou moins réguliers, les spécialistes de l'enseignement supérieur à se rencontrer et même à travailler ensemble sur le cas français. En 2001, cette communauté de recherche a donné naissance au Réseau d'Étude sur l'Enseignement Supérieur (RESUP), actuellement dirigé par Christine Musselin (Sciences Po Paris).

En 1969, Serge Vassal envisage pour la première fois l'existence d'une géographie de l'enseignement supérieur en France (Vassal, 1969). Après avoir réalisé plusieurs travaux sur la question, il souscrit à leur association à la géographie sociale défendue par Robert Hérin (Vassal, 1993) à travers son projet de géographie scolaire (Hérin cité par Frémont, 1986). Pour Vassal, la géographie française de l'enseignement s'inscrit aussi dans le prolongement de la géographie de l'instruction apparue dans les années 1970 en Allemagne et au Royaume-Uni (Vassal, 1978, 1982). En 1988, il publie sa thèse d'État sur la géographie des universités européennes (Vassal, 1988). Il y montre qu'avec la massification scolaire, des questions spécifiques d'aménagement du territoire se posent pour les universités. Plusieurs types et logiques d'implantations universitaires sont distinguées selon le rapport qu'entretiennent les universités avec leur environnement urbain et régional. En 1993 (*op. cit.*), Vassal insiste sur le rôle de la DATAR pour que la France rattrape son retard en géographie de l'enseignement. Et pour cause, c'est au rythme des commandes de l'État que cette géographie se développe, surtout à partir du plan Université 2000 (U2000) adopté en 1990. D'après la revue de littérature « Enseignement supérieur, recherche et collectivités locales » : « alors que les géographes et les sociologues avaient été peu associés aux développements universitaires des années 1960, ils furent en première ligne pour la mise en place du Plan Université 2000 » (Lebeau & Vadelorge, 2012). Le plan U2000 aurait ainsi permis l'émergence des recherches sur la « territorialisation » de l'enseignement supérieur et de la recherche. D'une part les établissements ont dû faire face à une « hausse rapide des effectifs qui les a conduit à augmenter leurs capacités d'accueil », et d'autre part les collectivités locales (à l'exception des départements) ont obtenu de nouvelles compétences en matière d'enseignement supérieur et ont été incitées à participer financièrement (par exemple *via* les plans États-régions) à la densification et à la modernisation du parc d'enseignement supérieur (*ibid.*). Dans la mouvance d'U2000,

sous l'égide de la DATAR et du Plan Urbain, trois angles d'attaque ont été privilégiés à partir de 1990, qui ont indirectement posé la question de la répartition géographique des activités de recherche : le fait universitaire en ville moyenne, l'aménagement de la carte universitaire et le développement régional. Dans le cadre de chacune de ces approches, nous allons voir dans quelle mesure les activités de recherche scientifique sont ou peuvent être prises en compte.

Comme le Plan U2000 donne lieu à l'ouverture de nouveaux sites d'enseignement supérieur en villes moyennes (parmi lesquels les « antennes universitaires »), leurs cas particuliers attirent l'attention. À Toulouse, un nombre considérable de chercheurs se sont penchés sur cette question et Daniel Filâtre, le premier. Sociologue et élève d'Alain Bourdin, il s'intéresse au rôle joué par les élus locaux dans les délocalisations universitaires. En 1994, il aborde le cas des villes de Midi-Pyrénées dans le premier livre collectif coordonné par Michel Grossetti sur la « territorialisation » de l'enseignement supérieur (Filâtre, 1994). La même année, il fait un état de la question à l'échelle nationale dans un ouvrage publié avec les sociologues François Dubet et François-Xavier Merrien, et les urbanistes André Sauvage et Agnès Vince (Dubet, Filâtre, Merrien, Sauvage, & Vince, 1994). Au milieu des années 1990, viennent les premiers travaux sur « les effets de ces implantations nouvelles » (Filâtre, Milard, & Viés, 1996), suivis au cours des années 2000 par des interrogations sur leur devenir (Grossetti & Filâtre, 2003). Entre temps, Filâtre engraine d'autres chercheurs de son laboratoire (le CERTOP pour Centre d'Étude et de Recherche Travail Organisation Pouvoir, ex ERMOPRES) sur ce créneau à commencer par Christelle Manifet, dont il dirige la thèse : « Gouverner par l'action : le cas des politiques universitaires de La Rochelle, Albi et Rodez » (Manifet, 2004). Spécialiste de l'action publique et sensible aux disparités territoriales, Catherine Soldano s'engage également sur cette voie. En 2015, elle coordonne avec Rachel Levy et Philippe Cuntigh l'ouvrage collectif le plus récent paru sur ce qu'il est désormais convenu d'appeler les « sites secondaires », ou « SUVM » pour « Sites Universitaires de villes moyennes » (Levy & Soldano, 2015). Filâtre en signe la postface, ce qui confirme l'appartenance du livre à une « lignée » (Filâtre, 2015). Pour Filâtre, « L'approche systémique revendiquée » dans ce livre permet d'aborder, sans les figer, les différentes dimensions du « fait universitaire » (*ibid*, p.154). De son point de vue, « L'université singulière n'existe pas et il serait plus précis de parler du fait universitaire en y englobant l'ensemble des actions et activités d'enseignement supérieur, de formation, de recherche, de transfert de technologie, de valorisation et de vie étudiante » et d'admettre « que ces activités se combinent sous des formes différentes suivant les contextes et les temps. » (*loc. cit.*). Ainsi, la question de la recherche scientifique en villes moyennes est très présente dans ce dernier ouvrage (6 chapitres sur 7), alors qu'elle était plutôt marginale au moment des délocalisations universi-

taires. À l'époque, il s'agissait surtout d'observer les stratégies mises en œuvre localement pour obtenir l'ouverture de formations universitaires, les conséquences locales en matière de démocratisation de l'enseignement supérieur et les enjeux locaux d'urbanisme associés aux nouvelles implantations (Séchet-Poisson & Peyron, 1994 ; Dubet *et al.*, op. cit. ; Lassave & Querrien, 1994 ; Peyron, 1996 ; Felouzis, 2001 ; Lebeau & Vadelorge, op. cit.). En fait, l'intérêt pour les activités de recherche menées dans les villes moyennes se concrétise à partir du moment où se pose la question de leur légitimité.

« Depuis 2007, la politique de l'enseignement supérieur change de cap en s'attachant davantage à la légitimité scientifique des sites universitaires. Les enjeux internationaux fondés sur la performance et l'innovation, remettent en question ce qui est lu aujourd'hui comme un émiettement, une dilution des forces et un manque de visibilité (loi LRU, 10 août 2007). Quelle place reste-t-il alors, pour des structures de recherche associées à des petits centres universitaires dont l'existence pourrait être compromise ? » s'interrogent Myrtille Moreau et Frédéric Tesson, deux géographes travaillant dans une université en pleine crise existentielle, l'Université de Pau et des Pays de l'Adour (Moreau & Tesson, 2011). Pour Cécile Crespy (Science Po Toulouse) et Soldano, ce « retournement de conjoncture » s'explique aussi par la stagnation, voire la baisse des effectifs enregistrée dans les SUVM : « Les SUVM doivent désormais justifier leur maintien et clarifier leurs missions » (Crespy & Soldano, 2015). Deux pistes se dégagent pour étudier les pratiques de recherche en villes moyennes : les entretiens qualitatifs avec des chercheurs affiliés aux SUVM (Losego, 2004 ; Tallec, op. cit. ; Levy & Bourbiaux, 2015) et l'analyse de la production scientifique des SUVM (Milard, 2012 ; Levy & Jégou, 2013 ; Levy, Sibertin-Blanc, & Jégou, 2013). Dans l'ensemble, ces travaux suggèrent qu'à l'encontre de nombreuses idées reçues, sur lesquelles nous reviendrons, il existe des « chercheurs stars » et de la « recherche d'excellence » en villes petites et moyennes ; et que la production et la visibilité scientifique de ces territoires est en constante progression depuis les années 1980.

Alors que le « fait universitaire » en ville moyenne a surtout été étudié par des sociologues, le thème de la « carte universitaire » est plutôt une affaire de géographes. Un travail remarquable de cartographie a accompagné le développement de cette thématique comme en témoignent les atlas publiés par le GIP RECLUS entre la fin des années 1980 et la fin des années 1990 (Brocard *et al.*, 1996 ; Saint-Julien, 1999). À l'occasion d'U2000 – comme pour le plan suivant à la fin des années 1990 (U3M pour « Université du Troisième Millénaire ») – l'aménagement de la « carte universitaire » est gouverné par la volonté de faire correspondre l'offre d'enseignement supérieur à la demande étudiante. Cela s'explique par la nécessité de gérer les effets de la massification scolaire (Frémont, 1990). Un numéro spécial de l'*Espace Géographique* paraît sur ce thème en 1990, occasion pour

certain contributeurs de regretter cette approche partielle de la carte universitaire. En particulier, Thérèse Saint-Julien et Roger Brunet s'accordent à dire que du fait de son association à la fonction recherche, l'enseignement supérieur ne devrait pas être traité comme un « service banal » de proximité à la population ce que, pourtant, les politiques successives de décentralisation ont eu tendance à faire (Saint-Julien, 1990 ; R. Brunet, 1990). Pour Saint-Julien, les « activités [de recherche] sont beaucoup plus tributaires des avantages d'agglomération, et des avantages de site et de situation, que de la distribution géographique de la population » et d'après elle, « cette sensibilité particulière des activités de recherche freine la tendance normale à la diffusion des établissements d'enseignement supérieur, du haut vers le bas de la hiérarchie urbaine » (Saint-Julien, *ibid.*). Chez Brunet, la recherche devrait « être considérée comme une activité productive » et suivre une logique de localisation qui ne soit pas celle d'un « service banal à la population » (Brunet, *ibid.*). D'après Brunet et Saint-Julien, il y aurait donc deux tendances contradictoires dans la répartition des établissements universitaires : celle associée à l'offre d'enseignement supérieur pousserait à la dispersion des établissements, ce qui ne serait pas le cas de celle associée aux activités de recherche. Pour répondre à la demande de formations, la tendance à la dispersion ou déconcentration des établissements a été dominante des années 1960 aux années 2000, ce qui justifie le fort intérêt qui a suivi pour la carte des formations associé à la problématique de la démocratisation de l'enseignement.

Dès la fin des années 1990, Myriam Baron constate que « l'équipement de formations supérieures est maintenant aussi répandu que l'équipement dans lequel s'achève la scolarité obligatoire » de sorte que « tous les grands pôles universitaires complets se situent à moins de deux heures de route des autres pôles de leur région ou des régions voisines » (Baron, 2010). Or, cette situation ne correspond pas à une parfaite démocratisation de l'offre d'enseignement supérieur puisque les types de formation ne sont pas répartis uniformément sur le territoire national (Figure 8). Dès lors, même en ne tenant pas compte de la répartition des activités de recherche, il n'est pas possible d'« assimiler l'enseignement supérieur à un service homogène » (Saint-Julien, *op. cit.*). Cette spécificité, que Baron a perçue et analysée dès sa thèse dirigée par Saint-Julien sur les Instituts universitaires de technologie (Baron, 1994) est essentielle pour légitimer l'existence d'une géographie de l'enseignement supérieur. Après sa thèse, Baron poursuit ses recherches sur la carte des formations et des établissements du supérieur et parvient à faire exister cette géographie en France, avec quelques collègues dont Patrice Caro (Caro, 1994, 2003, 2006), Cathy Perret (Baron & Perret, 2006), Sandrine Berroir (Baron & Berroir, 2007) ou encore Charlène Le Neindre et Christine Zanin (Baron, Le Neindre, & Zanin, 2008). Mais, contrairement à ce que l'on pourrait penser, ce mouvement de recherche n'a pas vu le jour avec le plan U2000, il a simplement été conforté par cette opération politique.

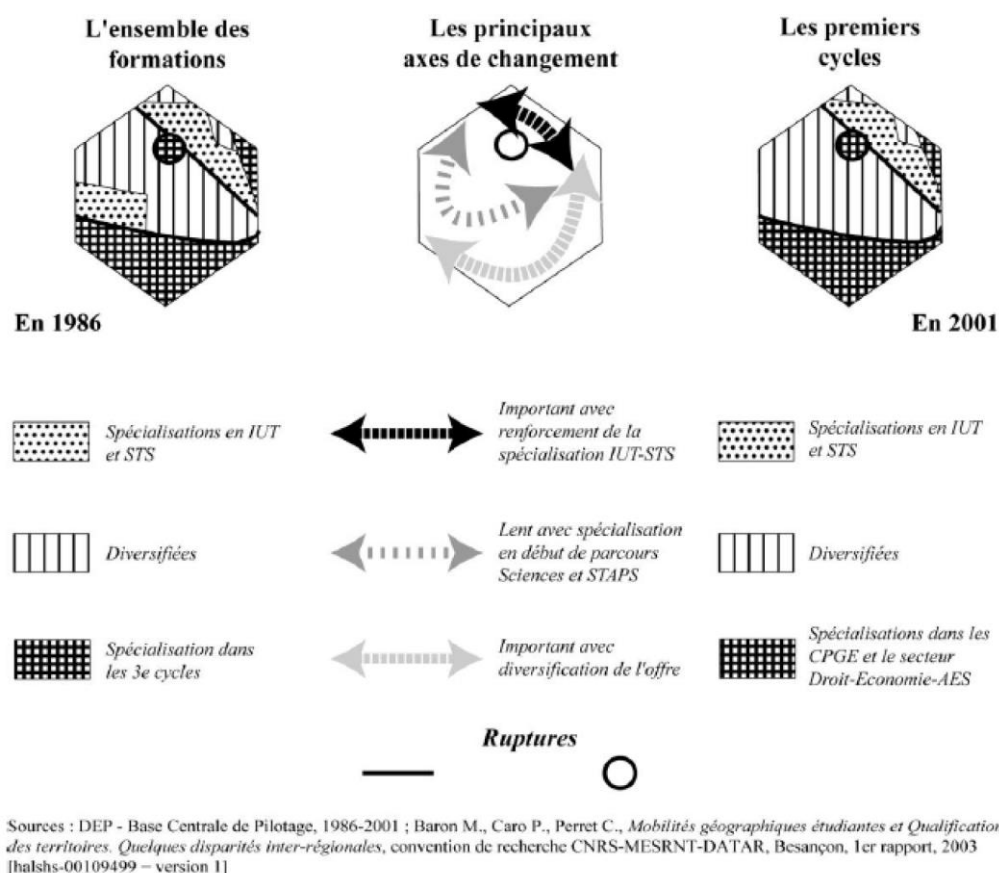


Figure 8 — Spécialisations universitaires « régionales » entre changements et reproductions (Baron, 2010)

Madeleine Brocard est à la fois l'une des premières géographes à s'être penchée sur les évolutions de la carte universitaire et l'une des seules à avoir pris le problème du point de vue de la recherche. Afin d'interroger le rôle supposé de l'activité scientifique pour le développement régional, elle a étudié la mise en place et les effets de la décentralisation des établissements de recherche à l'échelle nationale. La tâche n'était pas simple face au manque de statistiques publiques et à l'abondance de croyances invérifiées sur l'inscription spatiale des activités scientifiques (qu'elle n'a pas manqué de dénoncer). Sa réflexion débute « en 1969, dans le cadre d'un doctorat de géographie, en réponse aux interrogations de la DGRST [Délégation Générale à la Recherche Scientifique] » (Brocard, 1991, p. 12). En 1978, elle soutient sa thèse d'État intitulée : *Recherche scientifique et développement régional en France* (Brocard, 1978). Dans ce travail, Brocard revient sur l'évolution des politiques scientifiques depuis la fin de la Seconde Guerre mondiale. Elle montre que si la croissance des activités scientifiques a été considérée depuis longtemps comme une condition du développement économique national, son rôle dans la planification au niveau local a été négligé. Concrètement, il a fallu attendre le V^e plan de développement économique et social pour que le niveau local soit pris en considération. Comme son étude résulte

d'une demande politique, Brocard prend comme unité d'analyse les « régions de programme » pour lesquelles elle dispose de données remontant à 1963 (*ibid.*, p. 3).

En même temps qu'elle réalise des entretiens auprès de chercheurs tout juste installés « en région » – qui, à l'exception de ceux occupant des fonctions administratives, se déclarent satisfaits de leurs nouvelles conditions de travail et faiblement handicapés par leur éloignement⁵ – elle remarque sur le terrain que le groupe « localisation » constitué au sein de la Commission de la Recherche rencontre des difficultés à convaincre les organisations scientifiques de se décentraliser (*op. cit.*, 1991, p. 10). Pas de doutes pourtant du côté de la DATAR, devenue CGET en 2014 (Commissariat général à l'égalité des territoires) : les activités de recherche doivent être envisagées comme un levier pour le développement régional. Dans cette ébullition, Brocard a conscience de l'exceptionnalité de son approche : « Dans la littérature géographique, il n'existe pas de publications consacrées spécifiquement à la recherche, autres que celles qui ont été menées à Rouen par quelques chercheurs. » (*op. cit.*, 1978, p. 2). Elle pense à Romuald Leclerc qui a soutenu sa thèse de 3^e cycle sur la décentralisation de la recherche en 1974 (R. Leclerc, 1974), mais aussi à Pierre-Henri Emangard qui s'est penché avec elle sur le développement des activités de recherche en Bretagne (Emangard & Brocard, 1976). Leclerc n'a pas poursuivi sa carrière académique et Emangard comme Brocard, tous les deux en poste en Normandie dans les années 1980, ont été amenés à explorer d'autres voies pour leur carrière comme la géographie des transports, qui est une branche plus légitimée de la discipline (Bakis, 1980).

Pour autant, Brocard ne se désintéresse jamais de la question et reste attentive à l'évolution de la carte de la recherche qu'elle remarque extrêmement dépendante de la carte de l'enseignement supérieur⁶. En 1991, elle publie *La science et les régions : Géoscopie de la France*, ouvrage de synthèse sur le sujet (Brocard, 1991). Elle y insiste sur le processus de « partage du savoir » entre les régions françaises, qui tient à la diffusion des moyens de formation supérieure et l'élévation générale du niveau de qualification qui en résulte

⁵ Les chercheurs étaient « presque toujours satisfaits par la décentralisation, qui a permis l'amélioration du cadre matériel » (Brocard, 1978, p. 148) mais le problème résidait dans l'inégalité d'accès aux échanges : « Il est évident que plus l'on trouve en un même lieu géographique les organismes consultatifs des entreprises scientifiques, les manifestations où se réalisent les échanges informels, plus les directeurs de laboratoires ayant la chance d'y être implantés auront accès à ces échanges. Et lorsque tout, ou presque, est à Paris, Paris devient la masse critique nécessaire, même si des scientifiques provinciaux assument, au prix d'une fatigue accrue, le handicap de la distance géographique. » (*ibid.*, p. 158-159).

⁶ « La répartition des chercheurs de l'Enseignement Supérieur, parce qu'elle est globalement liée à celle des étudiants (coefficient de corrélation de 0,995 !) se calque de plus en plus sur le niveau supérieur de l'armature urbaine française. » (Brocard, 1991, p. 73) ; « Bon an, mal an, bon gré, mal gré, la géographie des laboratoires du CNRS se calque progressivement sur celle des Universités, elle-même liée à la répartition des étudiants. » (*ibid.*, p. 75).

(*op. cit.*, 1991, p. 109). Pour elle et malgré ces évolutions, « la rupture négociée du système hiérarchique et spatial liant Paris à la Province » à travers la décentralisation politique (*ibid.*, p. 257-258) est indispensable pour atteindre l'objectif recherché : le « développement économique [des régions] » (*loc. cit.*). C'est grâce au fait qu'elle aborde la recherche sous l'angle du développement régional (comme plus tard Patrice Caro en 1994 et Bertrand Moro en 2006 dans leurs thèses portant sur des contextes régionaux spécifiques), qu'elle est confrontée à la nécessité de mettre en regard les travaux de géographie économique sur les technopoles (abordés en Chapitre 2.2) et les travaux de géographie sociale sur l'enseignement supérieur. Cette démarche est essentielle parce qu'elle permet de détacher la question de la recherche de la seule question du transfert technologique, qui préoccupe tant les géographes de l'innovation. En matière d'aménagement du territoire, cela suppose de réfléchir aux moyens de concilier plusieurs finalités : la demande étudiante (formation), les attentes des industriels (transfert technologique), voire la soif d'apprendre des citoyens ou demande sociale (vulgarisation et patrimoine scientifique). Cela n'a rien d'évident parce que traditionnellement dans le domaine de l'action publique, formation et recherche sont des activités envisagées séparément, notamment à l'OCDE. On retrouve aussi ce clivage dans des organismes comme l'Institut d'Aménagement et d'Urbanisme de la Région Ile-de-France (IAU-IDF). Dans ce grand bureau d'études à visée opérationnelle, le département « Démographie, habitat, équipement et gestion locale » s'occupe des questions de formation (de Berny-Riche & Maisonobe, 2009 ; de Berny-Riche, 2013), et le département « Économie et développement local » aborde la question de la recherche et de l'innovation (Soulard & Guery, 2004 ; Soulard, 2010 ; Perrin & Soulard, 2014). Tout semble indiquer que les activités de recherche sont au service du développement économique des territoires quand les activités de formation contribuent à leur développement social.

Le parti pris de Brocard d'envisager la recherche en considérant son lien à la formation est donc tout à fait intéressant. En revanche, le fait d'aborder l'enseignement supérieur et la recherche seulement par le biais du développement régional (même si elle l'envisage à la fois du point de vue social et économique) est discutable parce qu'il peut amener à négliger les influences d'autres finalités qui dépassent le cadre régional ou qui regardent des niveaux inférieurs, comme le niveau urbain. On pense surtout à une finalité propre à la recherche scientifique, qualifiée par Michel Grossetti de « production d'universaux » (Grossetti, 1994). Roger Brunet est de cet avis, considérant qu'il faut se méfier de cet engouement pour le développement régional s'il implique la transformation des universités « en fournisseurs de services au détriment de leur production de savoirs et compétence » (*op. cit.*, 1990). Surtout, à force de prendre le problème dans ce sens-là, c'est-à-dire sous l'angle du rôle de la recherche pour le territoire, on en oublie le pro-

blème inverse, déjà évoqué et sur lequel nous reviendrons, celui de l'effet du territoire sur la recherche, voire sur les contenus scientifiques eux-mêmes. Pour échapper aux limites d'une vision trop utilitariste de la recherche faisant peu de place au concept de science et aux enjeux internes aux spécialités scientifiques, l'approche socio-historique des « pôles scientifiques » fait figure de troisième voie (Rollet, 2009). Parce qu'elle est moins sensible aux finalités immédiates prêtées à la science, l'approche socio-historique est à même d'expliquer la configuration actuelle de la carte d'enseignement supérieur et de recherche en la réinscrivant dans le temps long.

2. La géographie de la recherche : quelle est la place des logiques scientifiques ?

« Les classements sont instrumentalisés... Il y a une politique de dénigrement, une politique de la peur. On stigmatise, puis on annonce une modernisation. Ils le font pour tous les secteurs. Ils utilisent des faits divers etc. Le classement de Shanghai tombe, et ça sert à justifier une grande réorganisation... » Isabelle This Saint-Jean, vice-présidente chargée de l'enseignement et la recherche à la Région Ile-de-France, extrait d'un entretien, Janvier 2011

2.1. Pôles scientifiques et gouvernance universitaire

Le développement d'une approche sociohistorique des « pôles scientifiques » français doit beaucoup au travail de Michel Grossetti. Son itinéraire au sein du campus du Mirail, qu'il retrace dans son Habilitation à Diriger des Recherches (Grossetti, 1997), l'a conduit à côtoyer de près et même à intégrer (pour quelques années à partir de 1989) le Centre Interdisciplinaire d'Études Urbaines (CIEU). Sa collaboration avec des membres du CIEU démarre en fin de thèse, au moment où son directeur Jean-Michel Berthelot l'invite à s'insérer aux réflexions menées localement dans le cadre du programme PIRTTEM (Programme Interdisciplinaire de Recherche sur le Travail, les Technologies, l'Emploi et les Modes de vie). C'est au cours d'une réunion de travail en 1985, qu'il attire l'attention de ses collègues sur les « dimensions locales de la production des qualifications », sujet qui sera abordé dans le programme par le biais d'un axe sur la « technopole » de Toulouse. Cette thématique est approfondie ensuite à travers plusieurs projets. L'un d'entre eux, financé par le Plan Urbain, à l'occasion d'un appel sur « l'Université et la

ville » au début des années 1990, porte sur la genèse du pôle d'enseignement supérieur toulousain (Grossetti, 1994). Ce travail confirme l'intérêt de Grossetti pour l'histoire ; intérêt qui l'amène à faire la connaissance d'André Grelon, avec lequel il s'associe en 1995 pour le projet PIR-Villes du CNRS « Villes et Institutions Scientifiques ». C'est l'occasion d'appliquer à plusieurs villes la démarche mise au point pour étudier le développement du pôle toulousain. En préambule au rapport final, les auteurs vantent l'originalité de leur approche :

« En formulant la question des rapports entre villes et institutions scientifiques, notre volonté est de contribuer à ouvrir un nouveau domaine d'investigation, au carrefour de l'histoire et de la sociologie des organisations scientifiques, des études urbaines et de l'analyse des systèmes locaux d'innovation. Nous avons la conviction que cette ouverture est nécessaire et que ce domaine est destiné à prendre de l'importance. Nous faisons en effet l'hypothèse que les rapports entre villes et institutions scientifiques sont décisifs tant pour les villes que pour la science. Si la science devient une force productive de plus en plus importante, alors les laboratoires et les instituts de formation contribueront au destin des villes, au même titre que les usines, et même avec plus de force. Si les chercheurs, les enseignants et les étudiants sont de plus en plus nombreux, alors ils doivent imprimer leurs marques aux villes dans lesquelles ils vivent. Réciproquement, les hommes et les organisations qui font la science sont, d'une façon ou d'une autre, ancrés dans leur ville et peuvent être affectés à un moment ou un autre par les incitations, les pressions, les opportunités, issues de la ville comme entité à la fois politique, économique et sociale. » (Grossetti *et al.*, 1996).

Pour saisir simultanément les rapports de la science non seulement avec l'industrie, mais aussi avec l'enseignement supérieur, la ville est présentée comme un niveau d'analyse idéal par ces auteurs. Ces histoires locales montrent que les « pôles » d'enseignement supérieur et de recherche ont chacun leur spécificité et que les différences observées « ne sont pas réductibles à des effets de structure urbaine (taille des villes, compositions des populations) ou de spécialités industrielles » (Grossetti, Milard, & Maisonneuve, à paraître).

Aujourd'hui, avoir connaissance de ces différences et connaître leurs origines est capital dans un contexte où il est question d'améliorer la lisibilité de la carte nationale d'enseignement supérieur et de recherche et d'inciter les différents acteurs du système à se fédérer pour adopter des stratégies communes à différentes échelles territoriales (locale, régionale, nationale, européenne). En effet, comme l'ont remarqué Jérôme Aust et Cécile Crespy quand ils sont allés interroger la mise en place sur le terrain en 2007-2008 d'un nouveau dispositif de fédération (voire de fusion) des établissements de formation et de recherche au niveau local (le PRES pour Pôle de Recherche et d'Enseignement Supé-

rieur, qui a changé de nom depuis) : « "L'insertion dans la concurrence internationale", la promotion "de l'excellence académique et scientifique", la nécessité "d'être visible à l'international" fonctionnent plus comme des aiguillons pour l'action que comme un cadre cognitif fournissant des solutions concrètes aux "problèmes" fortement différenciés d'un site à l'autre, de la coopération et du regroupement universitaire. » (Aust & Crespy, 2009). Au cours des années 1990, Michel Grossetti et ses collaborateurs utilisaient pourtant la notion de pôle sans vouloir la doter de connotation politique ou performative. Cette notion avait l'avantage de ne pas les rendre prisonniers d'un échelon territorial, celui auquel il faudrait agir (la ville ou la région), et de ne pas les limiter à un seul type d'institution, celui autour duquel tout devrait graviter (l'université). Mais, la notion était déjà chargée de représentations pour l'action. D'après Laurent Rollet, elle aurait « pris de l'importance dans le vocabulaire technocratique à partir du mouvement de modernisation et de massification de l'enseignement supérieur des années 1960-1970 » (Rollet, 2009). Ainsi, en même temps que Michel Grossetti et ses collaborateurs retraçaient la genèse de plusieurs systèmes locaux de formation, recherche et innovation qu'ils qualifiaient parfois de « pôles », les spécialistes de l'innovation et des Systèmes Productifs Locaux (SPL) d'abord préoccupés par les qualifications, ont de plus en plus intégré les dimensions formation et recherche dans leurs analyses (la production des qualifications) ; tant et si bien que ces trois dimensions en sont venues à être considérées sur le terrain de l'action publique comme le socle de la croissance de demain.

Dans le premier Schéma Régional conforme aux directives de la Loi de 2013 sur l'Enseignement Supérieur et la Recherche (ESR) en France, l'enjeu est de lier les trois dimensions ou pointes du « triangle de la connaissance » (formation, recherche et innovation) (Région Pays de la Loire, 2013, p. 27). D'une manière générale, les acteurs locaux de l'enseignement supérieur et de la recherche (les collectivités territoriales et les établissements) sont incités à développer des « stratégies territoriales » (pour les premiers) et des « politiques de site » (pour les seconds) qui intègrent la question de l'innovation et de son lien à la recherche. Cette intégration est légitimée par la nécessité de mobiliser des soutiens financiers privés pour financer la recherche. L'approche stratégique est aussi justifiée par l'impression d'un émiettement des établissements d'enseignement supérieur et de recherche – lié à la diffusion spatiale de l'offre d'enseignement supérieur – qui a gagné du terrain à la fin des années 1990. En particulier, l'enthousiasme associé à la démocratisation de l'enseignement supérieur a diminué parce que les effectifs universitaires se sont stabilisés, voire ont diminué dans les facultés des sciences. Comme le pays est entré dans une nouvelle phase de réformes qui prévoit d'améliorer la gestion des dépenses du budget de l'enseignement supérieur et la recherche, il s'agit d'inciter les acteurs locaux à se rassembler autour de stratégies communes.

C'est ce qui explique qu'une grande partie des chercheurs abordant la « territorialisation » de l'enseignement supérieur et de la recherche intègrent désormais la problématique de la « gouvernance » universitaire à leurs travaux, et ce faisant analysent les conséquences de l'application d'une logique de marché aux universités (Darréon & Filâtre, 2003). Le problème majeur soulevé par cette littérature porte sur la tension entre la différenciation des établissements à travers des mécanismes de hiérarchisation tels que les classements et les évaluations, et la normalisation ou la standardisation des établissements nécessaire à leur intégration dans la compétition, autrement dit à leur participation au marché de l'enseignement supérieur et de la recherche. Les contradictions relevées dans les politiques et les discours associés à cette marchandisation viennent du fait que l'enseignement, la recherche et l'innovation ne s'appuient pas sur les mêmes « valeurs » – si bien que l'on pourrait parler de plusieurs marchés : celui des qualifications pour l'emploi, celui du savoir pour la science, et celui de l'innovation pour la croissance économique – mais aussi que les enjeux peuvent varier d'une échelle territoriale à l'autre. Ainsi, dès 2004, Christine Musselin remarque qu'au niveau européen, le processus de Bologne lancé en 1999, associé à la volonté de construire un Espace Européen de l'Enseignement Supérieur (EEES), et le processus de Ljubljana lancé en 2007, associé à la volonté de construire un Espace Européen de la Recherche (EER), encouragent deux orientations disjointes : l'homogénéisation des filières et des diplômes, donc des établissements d'une part ; la concentration des moyens de la recherche dans certains établissements capables de s'imposer sur la scène internationale, donc la différenciation des établissements d'autre part (Musselin, 2004).

Dans les travaux récents sur ces questions, dont nous avons pris connaissance pour la première fois à l'occasion du colloque « L'enseignement supérieur et la recherche en réformes » organisé par le RESUP en janvier 2011, l'inquiétude se porte sur le développement d'un marché à deux vitesses comme celui des États-Unis et du Royaume-Uni. Dans ce dernier pays, la gouvernance de la recherche repose depuis la fin des années 1980 sur les principes du « Nouveau Management Public » (NPM) (Rebora & Turri, 2013). Ne s'appliquant qu'à la recherche, l'exercice d'évaluation des universités est, là-bas, directement connecté à la répartition des moyens financiers, et entretient la différenciation entre deux systèmes : le système des universités de recherche (« centrées sur l'enseignement en master et doctorat ») et celui des universités d'enseignement (« gérant surtout l'enseignement de premier cycle ») (Louvel & Lange, 2010). Aussi, bien qu'une telle distinction ne soit pas faite par le système français d'évaluation, le principe de la concentration géographique des moyens de la recherche est, depuis longtemps (au moins depuis la III^e République à travers la vision de Louis Liard, voir par exemple Baron, 2012), une constante des politiques françaises d'aménagement de la carte universitaire.

Au moment du plan Université 2000, cette idée est aussi associée à une vision européenne de la recherche puisqu'il est prévu de faire émerger plusieurs grands « pôles européens » (Frémont 1990). Madeleine Brocard remarque même : « la notion de "pôle d'excellence scientifique" revient régulièrement dans les discours concernant la recherche publique, puisqu'il s'agit de répartir les moyens de l'État. Elle s'appuie sur l'idée qu'il existe des effets de seuil : au-delà d'un certain seuil quantitatif de chercheurs dans une discipline donnée, la concentration de matière grise et d'équipements déclencherait l'étincelle. Cela n'a jamais été prouvé (...). » (Brocard, 1991, p. 49). S'appuyant sur cette croyance, les opérations associées au Grand Emprunt à partir de 2010 ont eu tendance à reconcentrer les moyens de la recherche. Des partenariats, équipes et laboratoires de recherche ont été sélectionnés pour bénéficier des « investissements d'avenir ». Orchestrée par le Commissariat général à l'investissement (CGI), cette procédure s'est inspirée de l'*Exzellenzinitiative* allemande (Louvel & Lange, *op. cit.*). Les lauréats ont obtenu des labels (idEX, labEX, équipEX etc.) assortis d'une somme d'argent pour mener à bien leur projet de recherche. Parallèlement, des investissements immobiliers ont été programmés dans le cadre du « Plan Campus » pour rénover les équipements d'un nombre limité de sites universitaires.

Mais, l'insistance sur quelques « pôles » qui auraient « la masse critique » nécessaire pour participer à un espace de recherche ou « marché » international pose problème dans les faits. D'une part, comme le suggérait déjà Brocard au début des années 1990 (*op. cit.*), il n'existe pas de preuve empirique démontrant l'existence d'une relation de causalité entre la concentration des chercheurs et la productivité des chercheurs⁷ (voir la démonstration en Partie 2. Chapitre 8.2.3). La preuve a été aussi bien faite pour le contexte français (Grossetti, Milard, & Losego, 2002 ; Grossetti, Losego, & Milard, 2009 ; Grossetti & Milard, 2011 ; Bouba-Olga & Grossetti, 2012), que pour d'autres pays comme le Royaume-Uni (Bonaccorsi & Daraio, 2005 ; Aston & Shutt, 2009 ; Adams & Thomson, 2011 ; Kenna & Berche, 2011 ; Ploux-Chillès & Heintz, 2012). Et d'autre part, il paraît difficile de priver des équipes de recherche de soutien public au prétexte qu'elles seraient mal localisées, surtout si leur localisation s'explique par l'objet des recherches scientifiques conduites par ces équipes. En effet, comme l'ont déjà montré plusieurs études notamment sur les petites villes, il y a, parmi les déterminants de la localisation des lieux de recherche, des facteurs intra-scientifiques qui tiennent entre autres à l'objet des recherches.

⁷ En effet, les exemples ne manquent pas de chercheurs qui, comme Albert Einstein, ont été capables de prouesses sans être quotidiennement entourés d'une « masse critique » de pairs dans leur environnement de travail immédiat, sans compter que le nombre de publications d'une ville ou d'une région est, en général, quasiment une fonction linéaire du nombre de chercheurs (voir Partie 2. Chapitre 8.2.3).

Sans compter qu'une telle réforme ne va jamais sans quelques concessions (l'histoire en est témoin) si bien qu'au moment où débute la deuxième vague d'« investissements d'avenir » en 2014, la stratégie France Europe 2020 donnant les grandes lignes de la politique nationale de recherche évoque les limites des « politiques d'excellence ». En particulier, elle intègre la nécessité de rééquilibrer les moyens : « D'importantes inégalités territoriales en matière de recherche ont toujours existé. En effet les politiques de recherche nationales et européennes, en encourageant une recherche académique au meilleur niveau international, favorisent naturellement la concentration des talents et la constitution de grands pôles de recherche dont l'attractivité croît avec le développement. Il est cependant indispensable de veiller à la cohérence globale du dispositif national de recherche et de soutenir partout l'émergence de projets scientifiques de qualité qui pourront être, demain, les points forts de notre dispositif. » (Ministère de l'Enseignement Supérieur et de la Recherche, 2013, p. 75). L'emploi de l'adverbe « naturellement » doit ici nous interpeller. Selon une rhétorique bien connue, ce discours invite à combattre les inégalités engendrées « naturellement » par la promotion de la « compétitivité » en assurant un certain degré de « cohésion territoriale ». La conciliation de ces deux principes, qui sont au cœur des politiques d'aménagement européennes, s'inspire d'une logique économique très simple : pour qu'une « concurrence loyale » puisse régner au sein de l'espace européen (en l'occurrence, de l'Espace Européen de la Recherche), les adversaires doivent pouvoir combattre à arme égale (Bourdin, 2014). Ici, l'intervention de l'État est présentée comme un moyen de pallier les distorsions du « marché » induisant un renforcement des inégalités territoriales.

Cette théorie s'appuie sur l'idée très répandue que, suivant une logique d'organisation typiquement économique, il y aurait un inévitable mouvement de concentration des forces de recherche dans quelques grands pôles. Pour s'en sortir, les régions aspirant à obtenir des Fonds européens dans le cadre de la politique de cohésion pour 2014-2020 sont incitées à développer une stratégie dite de « spécialisation intelligente » (*smart specialisation*), qui demande de sélectionner quelques secteurs technologiques clefs ou « avantages comparatifs » dans lesquelles leur « écosystème local » aurait intérêt à se spécialiser. Pour que cela fonctionne, les régions sont appelées à fédérer leurs acteurs locaux de l'enseignement supérieur, de la recherche et de l'innovation. La philosophie de cette approche est qu'au lieu de « saupoudrer » les moyens (y compris les moyens visant à réduire les disparités territoriales), mieux vaut concentrer les investissements publics sur des lieux d'excellence et des domaines d'excellence. Toujours selon cette philosophie, puisque peu de sites ont la dimension ou « masse critique » suffisante pour cumuler plusieurs domaines d'excellence à la fois, et à ce titre jouer le rôle de pièces charnières dans l'échiquier territorial, alors la majorité a intérêt à se spécialiser dans quelques niches stra-

tégiques si bien qu'au bout du compte, tout le monde aura sa place. Comme l'expliquent Dominique Foray et Xavier Goenaga, qui sont les promoteurs de cette stratégie, « si vous êtes petit (...), alors il faut que vous soyez plus intelligent. »⁸ (Foray & Goenaga, 2013). L'idée sous-jacente est donc bien qu'en matière de recherche la taille du « pôle » détermine sa polyvalence.

L'idée implicite, que seuls les pôles polyvalents puissent avoir une vocation européenne ou mondiale, et qu'à moins d'être connectées au grand pôle de leurs régions, les équipes des petits sites risqueraient d'être marginalisées, présuppose que les activités de recherche obéissent à une logique d'organisation spatiale de type *hubs and spoke*. En effet, cette idée implique l'existence d'un système d'échange scientifique hiérarchisé où les plus grands pôles participent aux échanges mondiaux quand les pôles « secondaires » communiquent, soit entre eux, soit avec le pôle dont ils sont le satellite. En interrogeant cette idéologie au niveau de la région Pays de la Loire et pour le cas de la chimie, Béatrice Milard a montré que la production et la visibilité scientifique des villes moyennes de cette région est en constante progression depuis les années 1980 ; que leur croissance progresse à un rythme supérieur à celui des plus grandes villes ; que les équipes de chimistes de ces villes n'ont pas nécessairement besoin de travailler avec les équipes des grandes villes voisines pour être connectées et visibles à « l'international » ; et enfin, qu'au lieu de s'hyperspécialiser, la chimie pratiquée dans ces villes moyennes a tendance à se diversifier avec le temps (Milard, 2012). Cette analyse suggère que l'organisation des réseaux de recherche obéit à des logiques scientifiques spécifiques plutôt qu'au principe de la concurrence économique, mais mérite d'être prolongée à d'autres espaces géographiques et généralisée à d'autres disciplines ou spécialités.

2.2. Villes, réseaux de villes et activités scientifiques

Prendre en compte les logiques internes aux activités scientifiques dans le cadre d'une géographie des activités humaines, et du point de vue des questionnements de la géographie urbaine en particulier, présenterait au moins deux intérêts :

- mieux comprendre et interpréter la carte des établissements de recherche et d'enseignement supérieur, ses spécificités et ses inerties ;
- saisir une dimension importante de la mise en réseau des territoires à plusieurs niveaux dont le niveau local ou urbain ; autrement dit, considérer la science comme un

⁸ L'original : « Small is not necessarily more beautiful in the information age. If you are small, you are not in a good position to benefit from returns to size and so you have to be smarter. » (Foray & Goenaga, 2013).

domaine d'activité favorisant les circulations et les échanges entre systèmes urbains à différentes portées.

Concernant l'inscription spatiale des activités scientifiques, deux idéologies opposées influencent les représentations : les vertus de la concentration spatiale des scientifiques d'une part ; le mythe du chercheur solitaire capable de faire d'immenses découvertes tout en étant retiré du monde ou dans une tour d'ivoire d'autre part.

La première que nous avons déjà évoquée consiste à assimiler les activités scientifiques à des activités qui, en raison de leur dimension innovante, auraient intérêt à se concentrer dans l'espace pour bénéficier des avantages liés à la proximité, au rang desquels la très décriée « masse critique ». De ce point de vue, les milieux concentrant un nombre important de chercheurs, d'industriels et même d'étudiants seraient les plus attractifs et les plus excellents sur le plan scientifique. Cette représentation est mobilisée pour justifier les politiques de concentration des moyens matériels et financiers de la recherche. En France, parmi les discours relevant de cette idéologie, celui de Laurent Davezies est exemplaire. Spécialiste français en économie des territoires, ce dernier suggère, dans son ouvrage *La République et ses territoires : La circulation invisible des richesses*, que la politique de « régionalisation de la recherche » serait « malheureuse » puisqu'elle se serait faite au détriment de la région francilienne, pourtant la seule région française capable de jouer dans la cour de Boston et San-Francisco (Davezies, 2008, p. 82-83). La recherche semble fonctionner en vase clos et de manière indifférenciée dans sa logique car il oppose « la recherche qui trouve », celle qui est concentrée sur le plateau de Saclay en région Ile-de-France, à « la recherche territorialement équitable », c'est-à-dire une recherche de second rang disséminée dans l'espace national à cause des politiques successives de décentralisation. En fait, Davezies possède la même vision centraliste de la recherche que celle détectée par Madeleine Brocard à la fin des années 1970 dans les discours opposés à la « régionalisation » (Brocard, 1978). Déjà au XIX^e siècle, beaucoup d'intellectuels ont la conviction que le rayonnement de la science française à l'international passe forcément par Paris (Strada, 1899). Aussi, l'idée que l'efficacité et le prestige de la recherche française seraient maximisés par la position des chercheurs dans la capitale est profondément enracinée dans « les bureaux parisiens » (Brocard, 1991, p. 12). Au cours des années 1990, cette vision se trouve confortée par les théories sur la métropolisation et les villes mondiales, qui sont privilégiées, on va le montrer, par les premiers géographes urbains abordant la science comme richesse.

Parallèlement, dans les discours visant à défendre le caractère universel de l'acte intellectuel, une conception quelque peu différente du rapport de la science à l'espace est défendue. Tel un moine dans son monastère qui dispose sur place de tout ce dont il a besoin pour se recueillir, le scientifique devrait tout au plus exercer dans un laboratoire

équipé et connecté aux différents réseaux d'infrastructure et de communication (électricité, transport, poste, internet et téléphone) sans avoir à se soucier du contexte immédiat qui l'entoure ; et cela parce qu'établir de pures vérités scientifiques est une activité exigeant un haut degré d'abstraction. Déjà, dans l'utopie de Francis Bacon, la connaissance est produite à l'écart de la collectivité au sein la « maison de Salomon » — qui aurait servi de modèle à la Royal Society de Londres (Bacon, 1627). Ce qui se passe dans la maison demeure confidentiel et ce ne sont que les inventions jugées « utiles » qui sont rendues publiques. De nos jours, le laboratoire est, selon Livingstone, emblématique de cette conception puisqu'on le suppose « *placeless* » (Livingstone, 2003, p. 3). Sa localisation n'a pas, en théorie, d'incidence sur son fonctionnement puisque ce qui s'y fait devrait pouvoir être fait de façon identique à des kilomètres de là dans un autre laboratoire lui ressemblant en tous points. À cet égard, il est prévu que la connaissance produite dans un laboratoire, ainsi que le protocole utilisé à cette fin, soient accessibles aux autres laboratoires, par-delà les âges et les frontières. Mais, comme le principe de la maison de Salomon l'indique bien, cette conception est plutôt réservée à la science dite « pure » par opposition à la science pouvant avoir des applications (Bacon, *op. cit.*). Cette dernière, en raison de sa moindre autonomie à l'égard de la société, serait vouée à produire un savoir éminemment plus corrompu par les intérêts extérieurs. Ainsi, dès le début du XX^e siècle le chimiste Henry le Chatelier, sensible aux rapprochements entre la science et l'industrie se désole : « C'est un préjugé commun à beaucoup de savants français de penser que la Science pure doit rejeter loin d'elle toute préoccupation pratique, qu'elle doit s'isoler de l'industrie, comme d'une promiscuité compromettante. » (Le Chatelier, 1925, p. 135). Aussi, dans un article intitulé « Le caractère international de la science et l'appartenance à une nation », le sociologue américain Norman Storer réserve son discours à la recherche fondamentale : « S'il est vrai que les phénomènes physiques fondamentaux sont universels et que les progrès de la science doivent en général suivre une voie unique, il s'ensuit que la position qu'occupe un savant dans l'espace (sinon dans le temps) n'a aucune incidence sur la mesure à laquelle il a la chance de contribuer à ces progrès — pourvu qu'il ait dûment accès aux informations relatives à l'état actuel des connaissances dans son domaine de spécialisation et qu'il dispose du matériel nécessaire pour entreprendre des recherches appropriées. » (Storer, 1970).

Dans un cas, le potentiel du chercheur à contribuer au progrès scientifique serait maximisé par son appartenance à un centre ; dans l'autre cas, du moment que le chercheur est bien informé et en mesure d'interagir à distance avec le centre, son potentiel de recherche est insensible à sa localisation. Dans les deux cas, cependant, le chercheur positionné au centre est avantagé. En effet, que l'on valorise l'influence d'autres intérêts comme les intérêts industriels ou que l'on se préoccupe de l'accès aux équipements, aux

informations, aux moyens de transport et de communications, les espaces « centraux » – autrement dit métropolitains – apparaissent comme privilégiés puisqu'ils remplissent par définition ces deux conditions (concentrer plusieurs types d'activités et disposer d'une bonne connectivité). Dès lors, il n'est pas étonnant que la géographie urbaine ait tendance à regarder la science comme une activité métropolitaine, mais nous pensons qu'il serait regrettable qu'elle se borne à la traiter comme telle. Qu'advient-il dans ce cadre des recherches menées sur des terrains exotiques ou bien dans des laboratoires isolés pour des raisons de secret militaire ? Quel sort doit-on réserver aux enseignants-chercheurs qui exercent dans les antennes universitaires des villes petites et moyennes ? De ce point de vue, la vision métropolitaine de l'activité scientifique comporte le risque d'entraver le champ d'investigation des géographes urbains.

Le travail pionnier du géographe Christian Matthiessen et ses collègues (Matthiessen & Schwarz, 1999 ; Matthiessen, Schwarz, & Soren, 2002 ; Matthiessen, Schwarz, & Find, 2002, 2010), s'inscrivant dans la continuité des réflexions amorcées par Åke Andersson (Å. E. Andersson & Persson, 1993), est représentatif de cette limite puisqu'ils cherchent à saisir l'évolution du système scientifique mondial uniquement à travers l'activité des villes publiant le plus d'articles scientifiques (Grossetti *et al.*, 2014). Non seulement cette vision est contraignante mais elle pourrait, se faisant « prophétie autoréalisatrice », enrayer le développement des territoires (Staszak, 2000). En effet, elle pourrait influencer ou du moins conforter les décideurs politiques, les entrepreneurs ou même les chercheurs lorsqu'ils doivent respectivement allouer des financements, investir quelque part, arbitrer entre plusieurs lieux de travail ; même si, heureusement, compte tenu du nombre de facteurs qui entrent en jeu dans ce genre de décisions, il est vraisemblable que les théories désincarnées ne les orientent qu'à la marge. En effet, à rebours des études de Richard Florida sur la grande mobilité des « créatifs » et leur attirance pour les métropoles (Florida, 2005a), plusieurs enquêtes révèlent que ces derniers (les scientifiques, les artistes...) ne sont pas des électrons libres (Martin-Brelot, Grossetti, Eckert, Gritsai, & Kovács, 2010). En cas de mobilité, le choix de résidence des « créatifs » s'explique davantage par leurs parcours de vie (parmi les déterminants majeurs, on trouve l'aspiration à retourner vers le lieu d'étude ou le lieu d'origine ou encore la nécessité de se plier aux contraintes du conjoint) que par le caractère métropolitain et le rang des lieux d'accueil potentiels. En plus, les activités de recherche sont très variées, de telle sorte que si un chercheur devait tenir compte d'une hiérarchie pour décider de son affectation, ce serait sans doute d'abord de la hiérarchie des laboratoires qui abritent sa spécialité de recherche. Pour autant, l'engouement pour les classements de villes entretient l'idée qu'un petit nombre de lieux compte à l'échelle mondiale, que leur rang est mesurable et que s'y concentre l'essentiel du dynamisme de la planète (Moriset, 1999).

Cette inclination pour les classements et la tendance à se concentrer uniquement sur les villes qui se trouvent en haut de la hiérarchie s'expliquent par le caractère très concentré d'une part importante des activités humaines visibles et génératrices de profits à l'échelle mondiale. Le constat de cette concentration tend à contredire les utopies, comme celle du « village global », qui sont inspirées par la perspective d'un accès généralisé aux moyens de transport, d'information et de communication et iraient davantage dans le sens d'une vision non métropolitaine de la recherche. En effet, l'intensification de la mondialisation, ou plus précisément du « sentiment d'appartenir à un monde dont les parties sont interdépendantes » (Arrault, 2007, p. 476), a nourri depuis les années 1960 une quantité de discours prophétiques sur la fin de la géographie, la fin des villes, l'abolition des distances ou encore l'homogénéisation des sociétés et des territoires. Parmi les nombreux écrits abondant en ce sens, figure le *bestseller* publié en 2005 par le journaliste américain Thomas L. Friedman *The world is flat : a brief history of the twenty-first century* (Friedman, 2005). Ce livre défend l'idée que, grâce à la mondialisation, l'espace mondial s'aplatit, c'est-à-dire qu'il s'homogénéise et se pacifie. Or, l'un des détracteurs les plus célèbres de ce manifeste n'est autre que Richard Florida. Interpellant directement Friedman, Florida publie, dans la presse américaine, un article intitulé « *The world is spiky* » (le monde est piquant) où il objecte que l'espace mondial est traversé par un renforcement des inégalités lié à l'inéluctable concentration des ressources et des richesses dans un nombre limité de lieux de la compétition internationale (Florida, 2005b). Il se fait ainsi le porte-parole de la théorie de la métropolisation, qui comme nous l'avons vu à travers l'évolution de la géographie économique (Chapitre 2.2), a servi de cadre d'analyse pour saisir le mouvement de désindustrialisation à partir des années 1980.

C'est qu'il existe, en géographie urbaine, une relation entre la théorie de la métropolisation et celle de la mondialisation. En effet, la géographie urbaine a été amenée, en même temps qu'elle prenait en charge le processus de désindustrialisation, à incorporer les questions économiques associées à la mondialisation – devenue le concept privilégié pour interpréter l'évolution des rapports de force, les échanges et les flux internationaux et transnationaux aux lendemains de la guerre froide. Parmi les auteurs français qui s'inscrivent dans ce courant, la métropolisation désigne le rôle joué dans l'économie mondiale – que Pierre Veltz qualifie d'« économie d'archipel » – par un ensemble de grandes villes qui concentrent le pouvoir, la population et les activités tertiaires et qui sont connectées entre elles par le biais d'infrastructures de transport et de communications : les métropoles (Veltz, 1996). Dans ce cadre, les métropoles sont les catalyseurs de la mondialisation, les intermédiaires des échanges et des flux mondiaux. Plusieurs groupes de recherche s'inscrivent dans ce champ de réflexion ce qui contribue au fait que le vocabulaire autour de la mondialisation soit si foisonnant, en témoigne le *Dictionnaire Critique de la*

Mondialisation dont la seconde édition est parue en 2012 (Ghorra-Gobin, 2012). Pour y voir plus clair dans ce champ, il est possible de distinguer l'approche française de l'approche anglo-américaine. Les deux approches, ceci dit, puisent leur héritage dans le travail pionnier du géographe Jean Gottmann au sujet de la *megapolis*, l'ensemble urbain qui s'étend de Boston à Washington sur la côte Est américaine, traduction spatiale d'un phénomène d'urbanisation intensive (Gottmann, 1957).

L'approche française se trouve entièrement formalisée et synthétisée pour la première fois dans l'ouvrage majeur *Mondes Nouveaux, projet de Géographie Universelle* coordonné par Roger Brunet et Olivier Dollfus (R. Brunet & Dollfus, 1990). Olivier Dollfus y forge le concept d'« Archipel Mégapolitain Mondial » à partir d'observations portant sur la Triade, c'est-à-dire les espaces dominants ou centraux du « système-monde » dans les années 1980-1990, localisés en Amérique du Nord, en Europe de l'Ouest et au Japon. Ces observations sont le fruit d'une démarche de géographie systémique inspirée par Fernand Braudel et Immanuel Wallerstein. Dans ce contexte, les spécialistes de la mondialisation comme Laurent Carroué ou Jacques Lévy prennent l'habitude d'évoquer le caractère mondialisé des activités scientifiques et leur dimension stratégique, mais pour étayer leur propos ils s'appuient rarement sur d'autres données empiriques que les données d'investissement dans la recherche et le développement (Carroué, 2007, p. 59 ; Lévy, 2008). Ainsi, malgré le consensus sur le fait que la science est un fait urbain mondialisé ayant son intérêt dans le cadre d'une géographie de la mondialisation, cette conviction reste largement inexploitée.

Bien que l'approche anglo-américaine, prenant « la ville mondiale » pour unité élémentaire d'analyse soit également inspirée par l'œuvre braudélienne, elle s'appuie surtout sur le travail du géographe britannique Peter Hall marqué par son passage au département d'études urbaines de l'Université de Berkeley en Californie pendant les années 1980. C'est sans doute là-bas qu'il rencontre Manuel Castells avec lequel il co-écrit un livre sur les technopoles : *Technopoles of the World* (Hall & Castells, 1994). Son approche insiste sur les systèmes d'échanges entre grandes villes ouvertes à l'international et donne lieu à la formation d'un réseau de recherche baptisé « *Globalization & World Cities* » (GaWC), actuellement piloté par des géographes britanniques de l'Université de Loughborough et américains du Metropolitan Institute de Virginia Tech (Ghorra-Gobin p. 412). Dans ce cadre, les auteurs qui comme la sociologue Saskia Sassen sont intéressés par la dimension économique des échanges parlent plus volontiers de « villes globales ». En fixant leur attention sur les lieux où l'activité humaine est la plus concentrée, ces spécialistes s'assurent d'une part que les opportunités d'interagir à distance grâce aux nouvelles technologies d'information et de communication ne suppriment pas les besoins de déplacements et de « face-à-face » mais participent à les renforcer ; et d'autre part, que le « ré-

trécissement du monde » ne se fait qu'au prix d'une concentration des activités au sein de villes « *hubs* » qui servent de relais aux échanges et aux flux. Au cours des années 2000, les échanges scientifiques sont abordés pour la première fois dans ce cadre, d'une part par Peter Taylor – un membre éminent du GaWC qui a co-signé plusieurs articles avec Ben Derrudder sur l'analyse du réseau des villes mondiales – et d'autre part, nous l'avons vu, par Christian Matthiessen et ses collègues. Sensible au rapprochement entre les *science studies* et la géographie historique, Taylor propose une analyse géohistorique de l'émergence de la science moderne, publiée dans la revue d'histoire *Minerva* (Taylor, Hoyler, & Evans, 2008). De leur côté, les géographes scandinaves traitent de la période contemporaine et étudient les réseaux de collaborations scientifiques entre « villes mondiales » (Matthiessen *et al.*, *op. cit.*).

Aussi fertile soit-elle, notamment parce qu'elle intègre la question des collaborations scientifiques interurbaines, cette approche métropolitaine de l'urbain présente trois grandes limites :

- premièrement, elle fait mine d'ignorer la possibilité d'un rééquilibrage et d'une réduction des inégalités entre les territoires. Or, sur le long terme (environ l'espace d'une génération), les services rares et les innovations se diffusent dans la hiérarchie urbaine, bien que suivant très imparfaitement une courbe logistique (Raynaud, 2010). Les équipements peuvent se banaliser et se répandre, modifiant les contraintes de localisation des professionnels qui y sont associés. Ainsi, les universités, qui sont des équipements urbains abritant des activités d'enseignement supérieur et de recherche, se sont largement répandues à partir du milieu du XX^e siècle et il s'en trouve désormais dans un grand nombre de pays et de milieux géographiques (Schofer & Meyer, 2005). En général, du moment que le décalage n'est pas trop grand entre les premières villes et les dernières et d'autant plus s'il existe un soutien politique en faveur de la réduction des « fractures » (Dupuy, 2007), les inégalités spatiales peuvent donc s'atténuer et les hiérarchies être bousculées ;

- deuxièmement, cette approche est insensible aux mouvements d'« urbanisation par le bas » ou de « mondialisation discrète » que des chercheurs comme Éric Denis et Olivier Pliez mettent en évidence sur leurs terrains d'investigation, respectivement l'Inde et le monde arabe (Pliez, 2007 ; Swerts, Pumain, & Denis, 2012). En effet, d'après François Moriconi-Ebrard, les analystes des tendances contemporaines mondiales ont tendance à négliger bon nombre de ces processus déterminants car les modèles explicatifs sur lesquels ils s'appuient sont souvent élaborés à partir du cas européen et de la logique marchande néolibérale (Moriconi-Ebrard, 2000 ; Moriconi-Ebrard, Denis, & Marius-Gnanou, 2010) ;

- troisièmement, en se concentrant uniquement sur les flux inter-métropolitains, l'approche par les « villes mondiales » laisse de côté les relations qu'entretiennent les têtes

de pont avec les villes petites et moyennes ainsi que la capacité des plus petites villes à échanger en s'affranchissant de la médiation des grandes (Bell & Jayne, 2009 ; Wolff, 1996).

Pour ces différentes raisons, il est progressivement devenu évident qu'un changement dans les présupposés était nécessaire en géographie urbaine⁹. Désormais, de plus en plus de villes peuvent aspirer au statut de « métropole » au sens où on l'entendait dans les années 1980. Les équipements de recherche tout comme d'autres équipements jugés rares dans les années 1980 et, à ce titre, distinctifs sont beaucoup plus répandus dans les années 2000. En effet, de plus en plus de villes disposent de laboratoires de recherche, d'aéroports ou encore de centres de congrès ou d'orchestres philharmoniques qui étaient les signes distinctifs d'alors (Reynaud, 1976 ; Labasse, 1982 ; Carroué, 1993) ; voire de gratte-ciel (Didelon, 2010). Bien sûr, il est intéressant de se pencher sur les nouveaux signes distinctifs : si bien que, dans le domaine des sciences par exemple, une hiérarchie urbaine peut être déduite de la distribution spatiale d'un domaine de recherche de pointe comme l'informatique dans les années 1980 (Grossetti & Mounier-Kuhn, 1995) ou les nanotechnologies dans les années 2000 (Comin, 2009) ; mais il importe d'être conscient du caractère éphémère des critères de hiérarchisation (exception faite de la dimension des villes en nombre d'habitants qui reste un facteur stable à plus long terme comme le signale Pumain, 1997). Ainsi, comme s'en amusent les détracteurs des classements de villes : compte tenu de l'hétérogénéité des lieux, tous ont leurs spécificités et sont susceptibles d'être les premiers quelque part (Éloire, 2010) ; et bien qu'il existe des effets de continuité (on parle de dépendance du sentier ou *path dependency*), l'histoire enseigne que les premiers servis sont toujours susceptibles d'être rattrapés ou dépassés. C'est en quelque sorte comme si les hiérarchies existaient pour être bousculées.

À ce titre, l'approche de Denise Pumain est exemplaire puisqu'elle replace les systèmes de villes dans leurs temporalités. En particulier, elle montre que les modèles prédictifs – comme celui expliquant la concentration des activités par la maximisation des économies d'agglomération – échouent à expliquer l'évolution de la répartition des activités humaines puisqu'ils ignorent le long terme : ils prétendent prévoir la géographie des activités au temps $t+1$ en s'appuyant uniquement sur son état au temps t (Pumain, 1997). En réinscrivant les dynamiques sur le long terme, Pumain démontre qu'il est nécessaire de ne pas seulement se focaliser sur le haut des hiérarchies à un instant t pour saisir un phé-

⁹ En particulier, c'est l'une des conclusions qui est ressortie des discussions entre les géographes Emmanuel Eveno, Denis Eckert et Eric Denis lors du séminaire L'Espace des Sciences Sociales organisé le 10 mars 2014 sur le thème « Catégorisation, Classements, Hiérarchie » à Toulouse. Ce point de vue est présent depuis longtemps parmi les membres de l'équipe du CIEU qui le développent à travers un axe de recherche sur les petites villes.

nomène d'expansion, tel que le développement d'une trame urbaine à l'échelle d'un territoire national ou continental (Bretagnolle & Pumain, 2010).

Plus largement, l'approche de Pumain et celle de Thérèse Saint-Julien participent depuis les années 1980 à alimenter tout un ensemble d'études quantitatives sur les systèmes de villes. Même si elles ne négligent pas la question de la mondialisation, les analyses de ce corpus sont plus souvent menées à l'échelle nationale et européenne. Comme on le verra dans la partie méthodologique, la définition de la ville et la délimitation des agglomérations urbaines occupent une place importante dans ce cadre. Au lieu de se concentrer essentiellement sur les plus grandes villes, une attention particulière est accordée aux hiérarchies dans leur ensemble, à la position des villes dans un système compte tenu de leur rang. Ces travaux – fondés sur des données recueillies empiriquement – montrent dans quelle mesure les positions des villes sont héritées du passé, mais aussi qu'elles sont sensibles aux aléas de l'époque contemporaine. Par exemple, une partie des recherches menées dans cette mouvance porte sur les villes dont le rang diminue à cause des effets de la désindustrialisation, les villes qui rétrécissent ou *shrinking cities* (Baron, Cunningham-Sabot, Grasland, Rivière, & Van Hamme, 2010). Lorsque l'échelle d'analyse est l'Europe – comme pour tout un pan des recherches de l'équipe P.A.R.I.S du laboratoire Géographie-Cité à Paris – l'objectif consiste plus spécifiquement à étudier l'intégration des villes dans un système européen (Cattan, Pumain, Rozenblat, & Saint-Julien, 1999).

Dans cette littérature, la science prend peu à peu sa place. Trois exemples révèlent une intégration progressive des systèmes de collaborations scientifiques dans ce cadre d'analyse :

- premièrement, citons la conférence « Villes et réseaux en Europe : quelles intégrations territoriales ? » organisée par le RIATE (Réseaux Interdisciplinaire pour l'Aménagement du Territoire en Europe) en 2005 qui a débouché sur un ouvrage coordonné par Nadine Cattan, *Cities and networks in Europe. A critical approach of polycentrism* (Cattan, 2007). Ce travail collectif comprend une étude de Francesca Silvia Rota sur les réseaux de recherche européens (Rota, 2007). Signe que la démarche n'est pas courante, Gabriel Dupuy évoque, dans une recension, l'originalité du chapitre sur les réseaux : « Il ne s'agit pas des réseaux classiques (transport, énergie, télécommunications,...) caractérisés par des infrastructures physiques mais des réseaux qui mettent en jeu des flux immatériels d'échanges et de coopération dans les domaines culturels, scientifiques, informationnels et dans un contexte mondialisé. » (Dupuy, 2008). Ces réseaux ne sont pas ceux dont s'occupe traditionnellement la discipline, ce qui ne veut pas dire qu'ils soient dépourvus d'intérêt mais, comme l'indique Dupuy, ils sont plus flexibles que les réseaux d'infrastructures.

- à partir de là, d'autres études suivent sur la même ligne, à commencer par la thèse déjà évoquée de Marie-Noëlle Comin « Réseaux de villes et réseaux d'innovation en Europe : structuration du système des villes européennes par les réseaux de recherche sur les technologies convergentes » soutenue en 2009 (*op. cit.*), qui donne lieu à un chapitre dans le 6^e volet de la série *Données Urbaines* (Comin, 2011) et s'accompagne d'une contribution du physicien Pablo Jensen, spécialisé dans l'analyse de la dynamique des systèmes complexes, sur l'évolution des co-signatures entre villes européennes (Jensen, Apolloni, & Rouquier, 2011) ;

- et enfin, plus près de nous, vient la thèse de Sylvain Cuyala sur l'histoire du réseau européen des géographes quantitativistes. On y découvre entre autres comment ce « mouvement » s'est structuré par une manifestation : l'ECTQG (*European Colloquium of Theoretical and Quantitative Geography*) (Cuyala, 2014). Cette étude fait aussi l'objet d'un article paru dans *Cybergeo* en 2013 qui présente l'originalité de s'inscrire, comme l'article de Taylor (*op. cit.*), dans le sillon de la géographie historique et d'apporter, en plus, une contribution à l'épistémologie de la géographie (Cuyala, 2013). Or, la géographie historique et l'histoire de la géographie sont les branches, comme on va le voir dans le chapitre suivant, à partir desquelles se structure une géographie de la science en Angleterre depuis le début des années 2000.

En effet, ces derniers domaines sont plus enclins à intégrer les questionnements propres au champ des *science studies*, qui œuvre depuis le milieu du XX^e siècle à saisir des phénomènes tels que l'organisation institutionnelle des activités scientifiques ou l'influence de l'environnement et de la culture des scientifiques sur le contenu de leurs recherches. Les spécialistes de ce domaine sont sensibles aux logiques internes à la pratique scientifique, aux valeurs guidant ces pratiques, aux variations intra-disciplinaires, aux mécanismes qui, comme la formation de collectifs, participent au développement scientifique. Ce faisant, la question essentielle de l'expansion du fait scientifique dans l'espace géographique ne leur a pas échappé. Aussi, même si l'idéal de l'autonomie et de l'universalité des pratiques scientifiques conduit parfois à occulter le rôle structurant de l'espace géographique, les spécialistes des sciences n'ont pas attendu que s'éveille, parmi les géographes, un intérêt pour la territorialisation des activités scientifiques, pour s'en préoccuper. Certains, comme Steven Shapin, ont même envisagé la nécessité de rediscuter cet idéal. D'autres ont mis à l'épreuve les frontières sur lesquelles il s'appuie comme la fameuse distinction entre sciences pures et appliquées. C'est ce qui explique qu'alors que peu de géographes aient contribué à l'avancée des *science studies*, des réflexions d'ordre géographique sont présentes dans ce domaine depuis ses origines.

CHAPITRE 4. ÉTAT DES SAVOIRS POUR UNE GÉOGRAPHIE DES SCIENCES

On l'a vu, dans le cadre de la géographie de l'innovation, la recherche n'est qu'une activité de médiation pour l'innovation et, dans le cadre de la géographie de l'enseignement supérieur, la recherche est envisagée indirectement à travers son rapport à la formation supérieure. Pour simplifier, tant qu'elle n'est appréhendée qu'à travers ces cadres théoriques, la science n'est mesurée que par le biais d'indicateurs indirects : d'une part, le nombre de brevets, d'entreprises de haute technologie et d'autre part, le nombre d'étudiants ou les investissements en recherche et développement. En analysant la recherche du point de vue de son caractère intrinsèquement créatif, en considérant notamment la production et les collaborations académiques des chercheurs, les géographes urbains ont emprunté une nouvelle voie pour traiter de « science » mais cette voie est encore trop étroite. En effet, l'intégration de la science dans la discipline ne devrait pas se limiter à nourrir des axes thématiques isolés (régionalisation et développement local, systèmes de villes et intégration européenne, mondialisation et métropolisation) puisque de par sa complexité et son autonomie recherchée vis-à-vis des autres faits sociaux, le fait scientifique est capable d'amener ses propres problématiques à la géographie, voire de bousculer les cadres établis.

Ainsi, le manifeste de Livingstone en faveur du développement d'une géographie des sciences s'appuie sur les traces et les pistes qui, dans la littérature existante en *science studies*, considèrent le rôle de l'espace (Livingstone, 2003). Même s'il reste parfois accessoire, l'espace géographique n'a jamais été absent des travaux de sociologie, histoire et anthropologie des sciences. Plus ou moins accentuée selon les périodes, la question de la spatialisation des activités scientifiques n'a, de ce point de vue, que trop tardé à mobiliser les géographes. Et, pour que la greffe prenne, il est indispensable qu'à la suite de Livingstone, les géographes intéressés par le fait scientifique s'aventurent dans la littérature existante en *science studies* afin de s'en nourrir mais aussi d'y contribuer.

1. La science dans l'espace et dans le temps avant les années 1990

« Du moment que vous avez accès à un téléphone, à une photocopieuse et à des fonds pour les congrès, vous êtes OK, vous êtes branché sur la seule université qui compte vraiment – le campus global. Un jeune homme pressé peut voir le monde en jouant à saute-mouton d'un congrès à l'autre. » (Lodge, 1984, p. 78-79)

1.1. De la communauté scientifique...

La géographie est d'abord abordée en sociologie des sciences par le biais de l'attention accordée aux institutions scientifiques, à leur fonctionnement, et à la circulation des modèles nationaux d'organisation. Le champion en sociologie des sciences de cette approche, que l'on peut qualifier de « structuro-fonctionnaliste » (Bourdieu, 2002b, p. 26) d'externe ou d'institutionnelle, est Joseph Ben-David. Son travail s'inscrit dans la continuité des travaux du père fondateur de la sociologie des sciences, Robert K. Merton.

Dans sa thèse publiée en 1938, Merton s'interroge sur les avancées scientifiques du XVII^e siècle et plus spécifiquement sur le dynamisme scientifique de l'Angleterre (Merton, 1938). Parmi les facteurs explicatifs qu'il retient, figure l'attitude de la société face au changement : « George Hakewell, Bacon, Hobbes, Boyle, Granville et beaucoup d'autres ont rejeté l'idée que le monde était immoral et voué à la destruction pour privilégier la croyance en un imminent et brillant futur. »¹⁰ (*ibid.*, p. 54). En plus de la disposition de la société face au changement, le lien entre développement scientifique et fortes densités de population est considéré. Merton pense que le besoin de science augmente dans les zones où la population se densifie, c'est-à-dire dans les zones en voie d'urbanisation. D'après lui, « une densité de population grandissante facilite les avancées scientifiques et techniques : premièrement, en stimulant de nouveaux besoins (ou en renforçant d'anciens) ce qui, plutôt que d'accélérer le taux d'inventions, a pour effet d'orienter les intérêts de recherche sur certains créneaux ; et deuxièmement, en revalorisant les activités innovantes puisque les débouchés économiques sont plus importants dans un large bassin de popula-

¹⁰ L'original : « George Hakewell, Bacon, Hobbes, Boyle, Granville and many others turned from the conviction that the world was degenerate and designed for destruction to the belief in an imminent and unequal brilliant future. » (Merton, 1938, p. 54).

tion. »¹¹ (*ibid.*, p. 575). En revanche, la densité de population agirait davantage comme une condition favorable au développement scientifique que comme une condition suffisante. Dans une société hostile au changement où le statut du savant n'est pas valorisé, la densité de population ne peut rien pour le développement scientifique. En effet, Merton considère que la densité de population ne suffit pas à maximiser les opportunités d'interactions sociales de la population qui sont, pour leur part, essentielles au développement scientifique :

« Un autre moyen par lequel on considère que l'accroissement de la densité de population peut accélérer le taux d'inventions, à savoir, en augmentant les possibilités d'interactions sociales, n'est pas exclusivement lié à l'accroissement de la densité, et pour cette raison, ce moyen doit être considéré comme un facteur indépendant de la concentration de la population. »¹² (*loc. cit.*). Cette proposition est fondamentale car elle évoque la méprise qu'il y a à considérer que des échanges fructueux résultent automatiquement de la concentration des individus. La concentration de la population est tout juste un facilitateur d'interactions. Bien avant le développement de l'économie géographique et le déferlement des questions de recherche tournant autour de la notion d'innovation dans les années 1980, Merton remarque : « Il est clair que le contact des esprits (dans certains contextes culturels) a tendance à stimuler l'observation et l'originalité. Les idées et les expériences qui seraient restées autrement purement personnelles peuvent être, au moyen de l'interaction, des sources d'innovations et de découvertes... »¹³ (*ibid.*, p. 576). Ce qui importe pour que se réalisent ces interactions utiles à la science, ce sont donc les opportunités de mise en contact des esprits. Ainsi, Merton présume que le développement scientifique du XVII^e siècle en Angleterre a été favorisé par l'établissement d'un système de diligences efficace sur le sol britannique mais aussi par une meilleure organisation du service

¹¹ L'original : « Increasing population density may thus facilitate the advance of technology and science primarily in two ways: first by evoking certain new needs (or emphasizing old ones) which has the effect primarily of directing inventive interest in certain channels rather than of accelerating the rate of invention; and second, by inducing a higher estimation of inventive activities because of the greater economic value of invention among large populations. » (Merton, *ibid.*, p. 575).

¹² L'original : « Another means through which the growing density of population is said to accelerate invention, namely, by increasing social interaction, does not necessarily arise from heightened density alone, and may hence be considered as a factor independent of the concentration of population. » (Merton, *loc. cit.*).

¹³ L'original : « Obviously the contact of mind with mind (within certain cultural contexts) tends to stimulate observation and originality. Ideas and experiences which would otherwise have remained strictly personal, may, through the medium of interaction, become elements of innovation and discovery. Observations may be made by one scientist for which he has no explanation. Were these observations not communicated to other investigators they would then have no significance for scientific development. But once they are submitted to others for explanation, once there is social interaction, there is a possibility (which is more probable, the more minds there are in contact) that these observations can be unified and systematized by a theory. » (Merton, *ibid.*, p. 576).

postal au niveau du Continent. Si ces aménagements devaient principalement profiter à la classe commerçante, les savants en ont tiré parti pour interagir davantage. Le « Grand Tour », à savoir le voyage initiatique dans les grands centres culturels européens, est même devenu le mode de socialisation de référence pour les jeunes intellectuels du XVII^e au XVIII^e siècle.

Enfin, l'idéal d'une communauté savante organisée au niveau du Continent s'est développé à cette période. L'image du « collège invisible », apparue dans les manifestes rosicruciens¹⁴ du début du XVII^e siècle, est reprise par un groupe de philosophes naturels organisé autour de Robert Boyle, actif dans la fondation de la Royal Society en 1660 (Yates, 2001, p. 236). Ensuite, d'autres sociétés savantes, conçues sur le modèle anglais, apparaissent un peu partout en Europe, participant à l'institutionnalisation d'un système scientifique de dimension européenne qualifié parfois de « République des Lettres » (Sigrist, 2013). Accompagnant la multiplication des savants et des cercles, les premières revues scientifiques voient le jour. Ces dispositifs institutionnels donnent progressivement lieu à une professionnalisation de l'activité, rendant « possible la reproduction de pratiques dans la longue durée », car comme le précise le sociologue et historien des sciences Yves Gingras : « sans institution, la science n'est qu'une forme de passe-temps privé » et dans ce cas, la « transmission des savoirs reste soumise aux aléas de l'histoire. » (Gingras, 2013, p. 30).

Pour résumer, Merton montre dans sa thèse que quand la confiance dans la science et les besoins de science sont stimulés, et lorsqu'un statut et une légitimité sont accordés aux savants alors il se trouve dans la population davantage d'individus attirés par la démarche scientifique. Si ces individus ont les moyens d'interagir alors les conditions sont réunies pour que s'institutionnalise un système de production scientifique. Le travail de Merton suggère que l'investissement de l'État dans la construction d'un tel système favorise son ancrage national. C'est d'ailleurs, parce qu'il est ouvert sur le monde mais se structure au niveau des États, que le système scientifique contemporain est décrit par Yves Gingras comme un « espace à la fois national (par ses ressources) et international (par ses produits) » (Gingras, 2002).

¹⁴ L'Ordre de la Rose-Croix est une société secrète qui fit beaucoup parler d'elle au début du XVII^e siècle suite à la publication de plusieurs manifestes qui firent polémiques dans toute l'Europe (Le Doeuff, 2000). Le doute plane sur son existence et sur ses membres : certains soupçonnent Francis Bacon d'y participer (ce qu'il nie) tandis que René Descartes tente vainement d'entrer en contact avec elle. C'est dans un manifeste rosicrucien qu'il est fait pour la première fois mention du collège invisible : « En ce qui concerne le Collège, je ne sais rien d'autre que ceci : Regarde autour de toi et prie sincèrement Dieu et tu le trouveras certainement. Les frères sont plus près de toi que tu ne le penses, qui que tu sois où que tu sois, près ou loin, bon ou mauvais, élevé ou bas, pauvre ou riche, près ou loin, et cependant ils n'ont pas le don de l'ubiquité, ni ne sont des artistes du diable, mais des Théosophes. » (Mögling, D. alias Schweighardt Constantiensem, 1618).

Afin de réfléchir à l'articulation entre les différents niveaux d'organisation, l'analyse proposée par Ben-David est à la fois subtile et pionnière (Ben-David, 1991). Au lieu de considérer l'organisation du système scientifique contemporain comme le seul résultat de politiques nationales, ce dernier parvient à repérer des mécanismes et des leviers locaux de développement déconnectés des pouvoirs centraux. D'après lui, la professionnalisation des carrières de recherche et le mouvement de spécialisation disciplinaire en résultant participent à la croissance de la science.

Il montre qu'au XIX^e siècle, l'institutionnalisation de la science en Allemagne, qui se manifeste par la création de laboratoires de recherche dans les universités, ne résulte pas d'une décision politique du centre ou de l'application d'une « conception philosophique de l'université » mais d'une série de décisions locales et désarticulées permises par le caractère décentralisé d'un système où les universités sont en compétition (Gad Freudenthal *in* Ben-David, *ibid.*, p. 97). Toutefois, cherchant à prouver la supériorité de l'organisation polycentrique qui caractérise les systèmes universitaires allemand et américain sur l'organisation centralisée de France et d'URSS, il néglige les pressions locales qui ont bel et bien participé à la reconstruction du système de recherche français après la Révolution. À rebours des analyses de Ben-David, l'historiographie récente montre que même dans un État centralisé comme celui de France et en dépit de l'autoritarisme de l'État napoléonien, la capacité des acteurs locaux à faire valoir leur point de vue a été décisive pour la dynamique du système : en témoigne l'exemple de la faculté des sciences à Toulouse et l'exemple des facultés de droit de province (Nye, 1975 ; Baron, Barrera, & Birck, 2015). Synthétisant ces analyses, Baron *et al.* remarquent que dans la France post-révolutionnaire, contrairement aux idées reçues, « l'implication des territoires est une constante du siècle : sociétés savantes, municipalités, industriels s'activent de toutes parts pour obtenir une faculté ou d'autres structures pré-universitaires qui seront essentielles pour l'avenir de l'enseignement supérieur français » et de la recherche (Baron *et al.*, *ibid.*). N'hésitant pas à attaquer les candidatures de leurs voisins, les élus français se concurrencent pour obtenir des facultés, ce qui suggère qu'en dépit du centralisme du gouvernement, le pouvoir de la base n'est pas inexistant.

Considérant les travaux de Ben-David et d'historiens des sciences plus contemporains, il est évident qu'il importe autant de tenir compte des logiques locales que nationales et transnationales pour comprendre la construction et le fonctionnement d'un système de recherche, et cela y compris dans un pays centralisé comme la France. Comme il tâche d'expliquer la domination successive de systèmes nationaux de recherche par des dysfonctionnements institutionnels, Ben-David dépasse les interprétations évolutionnistes simplistes. Dans cette perspective, il accepte l'idée, essentielle pour sa théorie, que plusieurs types d'institutions scientifiques (académies, universités, centres de recherche fon-

damentale ou appliquée) peuvent co-exister mais il n'envisage pas que plusieurs systèmes nationaux et modèles d'organisation puissent faire autorité en même temps. Sa théorie s'appuie sur l'idée que le centre de la « science mondiale » se déplace dans le temps : d'abord en Angleterre, il dérive vers la France avant d'atteindre l'Allemagne, qui sera détrônée par les États-Unis au XX^e siècle. Puisque Ben-David écrit pendant la guerre froide et qu'il est acquis au libéralisme du fait de sa trajectoire personnelle, il est certain de la supériorité du modèle américain et ne conçoit pas que la science soviétique puisse proposer un modèle alternatif influent. En s'appuyant sur ses observations empiriques des années 1950 aux années 1970, il déclare : « Pour le meilleur ou pour le pire, le système universitaire américain avec ses fonctions hétérogènes et sa permanente expansion, est maintenant le système le plus influent de l'enseignement supérieur et de la recherche » et il ajoute : « cette influence se diffuse partiellement par le biais des contacts internationaux croissants entre travailleurs académiques qui convergent désormais en Amérique comme ils eurent l'habitude de le faire par le passé en Allemagne. »¹⁵ (Ben-David, *op. cit.*, p. 151).

Parce qu'elle tient aussi compte des stratégies locales, l'approche de Ben-David se garde d'être diffusionniste à l'excès et c'est, à notre avis, ce qui la rend si pertinente pour les lecteurs contemporains. D'après lui, il existe une diffusion des modèles d'organisation de la science, mais l'appropriation des modèles varie suivant les contextes d'importation. Puisque sa théorie sociologique s'intéresse aux variables institutionnelles qui participent à l'émergence de la science contemporaine, elle présente la particularité de s'appuyer sur des sources historiques, ce qui la rapproche de la thèse de Merton et lui donne de l'audience en histoire sociale des sciences, mais c'est aussi ce qui l'éloigne de la sociologie des sciences, qui devient dominante dans les années 1970. Bien qu'il soit rigoureux et original, le projet de Ben-David est idéologiquement biaisé en faveur du modèle libéral américain d'où son décalage avec la mentalité rebelle et post-positiviste de la nouvelle génération de sociologues des années 1970 (*ibid.*, p. 23). En effet, le projet de Ben-David sert la conviction, qu'il partage avec la génération de Merton à la suite de Michael Polanyi, que le développement scientifique ne peut s'épanouir qu'en démocratie ou plus exactement dans « un climat de grande liberté intellectuelle » (Ben-David, 1970).

Cette vision, Michael Polanyi l'a défendue en opposition à celle des « humanistes scientifiques », un cercle britannique auquel ont appartenu Joseph Needham et John D. Bernal dans l'entre-deux-guerres. Séduits par l'idéologie communiste, ces derniers militaient en faveur d'une organisation planifiée de l'activité scientifique. Primordiale à leurs

¹⁵ L'original : « For better or worse, the american university system with its constantly expanding and heterogeneous functions, is now the most influential system of higher education and research » et il ajoute : « this influence spreads partly through the increasing international contacts of academic workers which converge nowadays on America as they used to converge in the past on Germany. » (Ben-David, 1991).

yeux, l'utilité sociale et économique de la science pour la société devait être garantie par l'État. Selon Ben-David, ce groupe entendait « fonder une sociologie systématique de la science » en s'inspirant « du marxisme et notamment de l'idée qu'ils se faisaient de la manière dont la science était planifiée et organisée comme élément de l'économie en URSS » (*ibid.*). Pour Michael Polanyi et au-delà pour Merton et Ben-David, non seulement cette vision mettait en danger l'intégrité de l'entreprise scientifique (qui, bien qu'encouragée par les besoins socio-économiques, devait conserver son indépendance) mais, d'un point de vue sociologique, cette vision avait le défaut de réduire la science à sa dimension utilitaire. Pour dépasser cette vision limitée, ils ont jugé préférable et plus heuristique de défendre l'existence (aussi utopique soit-elle) de « la communauté scientifique ».

Cette terminologie désigne l'ensemble des scientifiques exerçant leur travail en se référant à des valeurs et des normes compatibles avec le partage d'un idéal démocratique et susceptibles de faire consensus en dépit des particularismes et des intérêts nationaux et locaux. Pour Merton, les valeurs et les normes sur lesquelles se fonde la communauté (il parle d'*ethos*) sont celles dont l'adoption par les savants européens à partir du XVII^e aurait permis l'avènement de la « science moderne », et c'est donc en s'appuyant sur ses connaissances historiques que Merton a mis en évidence les normes de la communauté : universalisme, communalisme, désintéressement et scepticisme organisé (Merton, 1973).

Entre temps, le travail de Merton s'est un peu démodé car l'historiographie et la notion de « science moderne » ont évolué. En histoire des sciences, à la suite des travaux menés au début du XX^e siècle par Alexandre Koyré, l'expression « science moderne » désignait la science s'étant développée dans le sillage de la révolution copernicienne ou « révolution scientifique » du XVII^e siècle. Sous l'influence de savants européens comme René Descartes ou Francis Bacon, la science se serait progressivement affranchie de ses tutelles, et surtout de l'autorité de l'Église. C'est à partir de là, qu'illustrant l'incompatibilité entre la raison et le dogme, la condamnation de Galilée fut érigée en symbole pour l'autonomie de la pratique scientifique. Détachée de la foi et des intérêts particuliers, l'établissement de vérités « universelles » et « objectives » serait devenu l'une des finalités premières de cette pratique. Et ce sont ces valeurs qui, d'après Merton, continuent d'orienter l'action des scientifiques au milieu du XX^e siècle.

Le problème est qu'alors que les historiens se sont longtemps accordés à considérer que l'avènement de ce régime de vérité tranchait avec les systèmes scientifiques antérieurs ou co-existants qualifiés de « traditionnels » (notamment par Joseph Needham), cette conviction est aujourd'hui mise en cause. En histoire sociale des sciences, le passage de la science de la Renaissance à la science moderne n'est, en particulier, plus perçu comme une rupture (pour une synthèse des différents points de vue historiographiques, voir Cook, 2008). Il reste que comme l'ont souligné Merton et Ben-David et plus récem-

ment Yves Gingras, il y a eu un mouvement d'institutionnalisation des pratiques scientifiques européennes à partir du XVII^e siècle, qui peut être perçu comme constitutif du système scientifique contemporain. Certains auteurs parlent même d'une première division du travail scientifique qui se serait mise en place au XVIII^e siècle dans l'Empire britannique : celle-ci « repose sur des "corps" de scientifiques postés à travers l'Empire pour observer les phénomènes et recueillir systématiquement données et matériel. Ceux-ci sont adressés en métropole à des savants stationnés dans des établissements spécialisés, qui se chargent de l'interprétation et de la conservation » (Shinn, Vellard, & Waast, 2010). Dès lors, sans qu'il soit forcément judicieux de parler de « révolution », la mise en place de systèmes nationaux de recherche mais surtout la professionnalisation et l'organisation des savants au niveau européen et mondial a participé à l'émergence d'un « champ scientifique » que Pierre Bourdieu a défini pour la première fois en 1976 sur la base d'une analyse des pratiques scientifiques contemporaines, des normes et valeurs auxquelles elles se réfèrent (Bourdieu, 1976).

1.2. ...aux communautés scientifiques

Dans les années 1970, l'attention des spécialistes des sciences, souhaitant dépasser l'approche fonctionnaliste, s'éloigne des institutions et de l'organisation sociale des scientifiques pour aller vers l'élaboration concrète des contenus et la résolution des controverses. Dans cette mouvance, il devient plus difficile d'adhérer à la vision d'une « communauté scientifique » unifiée autour d'une même finalité (la vérité) et capable de l'atteindre par voie de consensus (Zuckerman, 1988). Rompant avec cette vision idéalisée, Bourdieu montre que les conflits et les rapports de domination sont structurants pour la pratique scientifique, mais, en bon « scientifique », exige que puisse triompher « l'idée vraie » en dépit des intérêts nationaux, locaux, disciplinaires ou même individuels qui orientent les actions et les modes de pensée des membres du « champ scientifique » (Bourdieu, 1976). En rupture plus forte encore, des spécialistes des sciences, déjà évoqués dans le Chapitre 1, prennent le parti épistémologique de considérer que « l'idée vraie » n'existe pas. Ces derniers bousculent le statut de la vérité scientifique en prônant une vision constructiviste, voire relativiste, de la connaissance scientifique. À cette fin, ils se lancent dans des études microsociologiques centrées sur les activités de laboratoire et les contenus scientifiques. Pour rediscuter la frontière entre « connaissance scientifique » et « connaissance commune », Bruno Latour et Steve Woolgar ont, par exemple, l'idée d'appliquer la démarche anthropologique à la vie de laboratoire (Latour & Woolgar, 1979). Ce faisant, la « nouvelle sociologie des sciences » s'écarte des méthodes privilégiées

par l'ancienne : le recours aux mesures bibliométriques et l'approche historique des institutions scientifiques (Dubois, 2001).

Alors que Ben-David ne percevait pas d'incompatibilité entre le travail de terrain microsociologique et ses propres travaux à caractère historique, la nouvelle génération fait tout son possible pour se démarquer de lui ainsi que du cercle mertonien (Gad Freudenthal *in* Ben-David, 1991). En dépit de cette querelle des anciens et des modernes, la revue *Social Studies of Science* (créée en 1970) publie des travaux issus des différents courants d'études sur la science, autrement dit : des travaux sur les institutions scientifiques y côtoient des études constructivistes sur les contenus scientifiques et des études bibliométriques sur les collaborations internationales. Mais avec le temps, la division des tâches s'intensifie entre les différentes sous-spécialités. Si l'on simplifie : les historiens et politistes s'occupent du travail sur les institutions de recherche ; les scientomètres se chargent des analyses bibliométriques ; et les sociologues des sciences se concentrent sur les analyses des pratiques et des contenus scientifiques. Parmi ces dernières, on distingue plus particulièrement les études qui considèrent que le niveau pertinent d'analyse de la « science en train de se faire » est celui des spécialités scientifiques, et les études qui préfèrent le laboratoire comme niveau d'analyse. On estime parfois que les secondes ont remplacé les premières pendant les années 1980 (Wray, 2005).

Victime de cette évolution et jugée inadéquate et trop abstraite pour décrire la réalité de l'entreprise scientifique concrète, l'expression « communauté scientifique » utilisée pour désigner simultanément tous les scientifiques œuvrant à la poursuite du progrès tous domaines de recherche et pays confondus, quitte momentanément la sociologie des sciences. Mais cette expression poursuit sa trajectoire parmi les spécialistes des pays en développement qui posent la question de la constitution de « communautés scientifiques nationales » au pluriel. Ces spécialistes, comme les membres français de l'équipe PED (Pays En Développement) créée en 1983 à l'ORSTOM (actuel IRD), ont une connaissance approfondie de leurs terrains d'étude. En s'appuyant sur leurs terrains, ils s'appliquent en premier lieu à rejeter la théorie diffusionniste selon laquelle la science moderne serait venue d'Europe pour s'implanter dans de nouveaux espaces nationaux et mobiliser des populations auparavant dépourvues de toutes formes de culture scientifique. Cette théorie généralement attribuée à Georges Basalla est une adaptation à la science de la théorie, déjà évoquée dans le Chapitre 2.2.1 (Encadré 1), des stades de développement de Rostow (Basalla, 1967).

Pour l'invalider, beaucoup de chercheurs y opposent la vision irénique de Joseph Needham, qui explique l'apparition de la science moderne en Europe par l'influence des sciences chinoises, arabes ou antiques. Selon ce dernier, c'est la rencontre des différents systèmes de pensées scientifiques préexistants qui débouche sur la science moderne. Il

propose une allégorie : « Il y a une vieille expression chinoise à propos des "Rivières qui vont faire la cour à la mer", et en effet nous pouvons très bien considérer les vieux courants de la science dans les différentes civilisations comme des rivières allant se jeter vers l'océan de la science moderne. La science moderne est en effet composée par les contributions des différents peuples du vieux monde, et chaque contribution se serait écoulée dans un flux continu vers l'océan, qu'elle soit de l'antiquité romaine ou grecque, issue du monde arabe ou encore des cultures chinoises et indiennes. » ¹⁶ (Needham, 1967). Dans cette vision « œcuménique » de l'entreprise scientifique, tous les peuples sont amenés à participer main dans la main à la progression de la science sans perdre leurs particularités.

Bien qu'opposées l'une à l'autre, la vision diffusionniste et la vision œcuménique sont toutes deux compatibles avec l'idée d'une « communauté scientifique internationale » et, en cela, se heurtent à la théorie de la dépendance qui se met en place à partir des années 1970. Pour Roy MacLeod ou encore Ian Inkster, il est important de distinguer le cas (qui sert de modèle à Bassala) des colonies de peuplement comme l'Australie ou le Canada où une portion de la population européenne est venue peupler de nouveaux territoires, de celui des colonies au sein desquelles des peuples autochtones sont assujettis par la force à une puissance européenne (MacLeod, 1980 ; Inkster, 1985 ; MacLeod, 1999). Dans ce second cas, l'expansion du système scientifique ne s'est pas faite sans heurts. En revenant sur l'histoire des violences et de la domination imposée aux peuples colonisés, il ressort que les conditions de participation au « champ scientifique » sont d'emblées inégales. Dans le contexte des décolonisations, les monographies de pays se multiplient qui rendent compte de la variété des conditions d'émergence des « communautés scientifiques nationales » dans les pays dits du « Tiers-Monde » (Vessuri, 2001 ; Paty, 2001). À une approche diffusionniste succède une approche plus culturaliste. Le rôle actif des élites locales dans la mise en place de systèmes nationaux de recherche, dans l'élaboration des savoirs et dans la lutte pour l'égalité est valorisé (Arvanitis, Grossetti, Raj, Renaud, & Thomas, 2008) ; si bien que l'idéal d'une « communauté internationale » unifiée est progressivement balayé. En particulier, en lieu et place de l'approche communautaire de la science ou des sciences, de nombreux chercheurs adoptent le modèle structuraliste décrivant un système mondial d'États-nations ou « système-monde » en expansion avec son centre et ses périphéries, son foyer de domination et ses dominés.

En effet, à partir du modèle des économies-monde de Fernand Braudel repris par Immanuel Wallerstein, Xavier Polanco propose l'expression « Science-Monde » (Polanco,

¹⁶ L'original : « There is an old Chinese expression about "the Rivers going to pay court to the Sea", and indeed one can well consider the older streams of science in the different civilisations like rivers flowing into the ocean of modern science. Modern science is indeed composed of contributions from all the peoples of the Old World, and each contribution has flowed continuously into it, whether from Greek and Roman antiquity, or from the Arabic world or from the cultures of China and of India. » (Needham, 1967)

1990). Le recours à cette terminologie s'accompagne d'une volonté de sortir du scientisme et de l'eurocentrisme dont sont imbibées les sciences sociales. D'après Patrick Petitjean, ces travers sont un héritage de l'idéologie impérialiste des sociétés européennes du XIX^e siècle. Membre de l'Union Internationale d'Histoire des Sciences à l'UNESCO¹⁷ et cofondateur de la commission "sciences et empires", Petitjean remarque que les savants n'ont pas toujours été aussi eurocentrés. Avant que la diffusion de la science européenne ou « civilisation » soit utilisée comme argument impérialiste par les élites : « les voyageurs européens reconnaissaient aux natifs leur participation à l'entreprise scientifique, et parfois même leur propre dépendance aux "informateurs" locaux. » (Petitjean, 2005). Ainsi, il semblerait que ce soit seulement vers la fin du XIX^e que les savants français aient commencé à revendiquer la suprématie de la raison scientifique cartésienne et à s'en servir d'argument pour justifier les colonisations. Pour Petitjean, cet état d'esprit est au fondement du modèle contemporain de la science et il est responsable de trois regrettables faits : « la négation des contributions non-européennes ; l'idée que la maîtrise de la nature est fondamentale pour le progrès social ; et l'idée que la civilisation "scientifique" d'Europe est un modèle qui devrait être suivi par les autres, un but à atteindre ». En fait : « ces deux tendances, le scientisme et l'eurocentrisme, sont mobilisées pour revendiquer une universalité "verticale", inscrite dans la nature. La "mission civilisatrice" fait alors figure de produit associé, dans lequel l'eurocentrisme et le scientisme se renforcent l'un, l'autre » (Petitjean, *ibid*).

L'analyse de Petitjean légitime la lecture critique des sciences sociales proposée par Wallerstein, l'un des pères de la théorie de la dépendance (Wallerstein, 1982). Cette lecture est intéressante dans la perspective de l'opération de décentrage qu'exige la fin des ethnocentrismes. D'après Wallerstein, nous l'avons déjà évoqué, les disciplines des sciences sociales ont été définies au XIX^e en s'appuyant sur les fondements idéologiques du libéralisme anglais. Puisque l'étendue des savoirs sur les sociétés humaines était encore limitée, les certitudes sur lesquelles s'appuyaient les premiers économistes, historiens et sociologues relevaient davantage de croyances que de faits établis. En particulier, l'État-nation et le capitalisme ont longtemps été appréhendés comme deux inventions occidentales susceptibles d'expliquer la supériorité de l'Occident sur l'Orient. L'État-nation était le moyen le plus évolué de faire société et le capitalisme était le mode de production le plus avancé que l'humanité n'ait jamais connu. L'historiographie offre de nos jours des visions beaucoup plus contrastées quant à l'origine et la nature de ces événements, mais également, on l'a évoqué, quant au rôle de l'Occident dans l'avènement de la science moderne. Cette transition fondamentale à la fois pour les sciences sociales et pour les *science studies* est évoquée par les historiens des sciences David W. Chambers et Richard Gillespie

¹⁷ Union jadis dirigée par Joseph Needham et Julian Huxley (dans les années 1950).

comme le passage en histoire des sciences d'une idéologie de l'inégalité entre les sociétés à une idéologie de la différence (Chambers & Gillespie, 2000).

Pour résumer, Wallerstein remarque que quand ils sont confrontés à plusieurs collectifs à la fois, les spécialistes en sciences sociales ont tendance à être : soit « universalistes » et chercher des invariants ; soit « culturalistes » et insister sur les différences, suggérant une sorte d'incommensurabilité mettant à mal la notion d'humanité au singulier. Entre l'option « universaliste » et l'option « séculariste », Wallerstein préconise une *via media* : « Nous devons apprendre à penser holistiquement et dialectiquement. Il faudrait dire "apprendre comment penser..." » (Wallerstein, *op. cit.*). Il ne s'agit pas d'être universaliste mais d'arriver à penser le monde dans sa totalité pour mieux en comprendre les parties. Il n'est pas question d'inférer un type de comportement humain pour en tirer des généralités sur l'humanité mais de rassembler des connaissances sur tous les types de comportement humain pour repérer ressemblances et dissemblances, complémentarités et incompatibilités. Puisque c'est à travers les organisations humaines que se passent les échanges et que l'humanité fait système, il convient d'étudier ces organisations et leurs dynamiques pour comprendre le monde. Or, et c'est l'un des points essentiels de son discours que nous voulons souligner ici, les États-nation ne sont qu'une forme parmi d'autres d'organisation humaine.

C'est pourquoi, Wallerstein regrette que notre appréhension du monde social soit limitée par notre accès à des données statistiques établies par et pour les États. Il explique : « Nous devons alors utiliser cette méthodologie pour inventer (j'emploie à dessein ce mot fort : inventer) les bases de nouvelles données. Celles que nous utilisons aujourd'hui (ou 98 % d'entre elles) ont été recueillies pendant 150-200 ans au sujet des États. Le mot même de "statistiques" est dérivé, et non pas par hasard, du mot "État" (*state*). Nous ne disposons pas de données sérieuses au sujet de l'économie capitaliste mondiale (sans parler des systèmes mondiaux antérieurs). Sans doute, la production de telles données présentera-t-elle de nombreux problèmes intrinsèques et extrinsèques. Mais l'ingéniosité méthodologique qui s'est manifestée au cours des 30 dernières années nous donne des raisons d'espérer. Elle a ouvert à la recherche quantitative différents champs de connaissance, réputés inattaquables à ce niveau, comme l'histoire médiévale. Nul doute qu'en y mettant assez d'énergie et d'intelligence nous puissions, dans 30 ans, recueillir autant de données concrètes sur le fonctionnement du système mondial moderne, comme système, que nous en avons aujourd'hui sur le fonctionnement des différents États. » (*ibid.*).

En prononçant cette conférence aux États-Unis en 1982, Wallerstein ne pouvait deviner sans doute combien les développements de l'informatique allaient venir faciliter l'accès, le traitement et la visualisation des données sur les hommes et leurs activités. Son

discours est d'autant plus prophétique qu'avec les crises économiques, les États ont été mis en cause par les chercheurs en sciences sociales et de nouveaux niveaux d'organisation ont attiré l'attention, notamment sur les métropoles comme nous l'avons déjà évoqué mais aussi sur les grandes aires macro-régionales comme l'Union Européenne. Cette évolution a été favorisée par le fait qu'avec la fin de la guerre froide, le marché a pris la place de la géopolitique pour comprendre l'ordre du monde (Laïdi, 1996). Si le travail de Wallerstein a certainement favorisé le développement de réflexions en termes systémiques ; il n'a pas fait disparaître le poids des constructions idéologiques en sciences sociales. Dans les années 1990, l'idéologie de la « mondialisation » tend progressivement à remplacer celle du « développement » pour saisir le monde.

2. Les sciences dans l'espace et dans le temps à partir des années 1990

« Les agents ne sont pas universels parce que leurs propriétés sont le produit de leur placement et de leur déplacement dans l'espace social, donc de l'histoire individuelle et collective. » (Bourdieu, 2000, p. 259 cité par Garcia-Parpet, 2014)

2.1. Du « tournant spatial » en science studies...

Dans les années 1990, de nouvelles questions affleurent concernant les rapports de la science au temps et à l'espace. Après le temps des controverses, vient celui des synthèses permettant d'articuler les apports des multiples études de cas réalisées pendant les années 1970 et 1980 en histoire, sociologie et anthropologie des sciences. L'entrée spatiale se révèle être un terrain propice à la réconciliation. Précisément, ce moment exige de s'appuyer sur de nouveaux modèles explicatifs, acceptables pour le plus grand nombre, c'est-à-dire ne pouvant être taxés ni d'occidentalisme, ni de positivisme. De plus, il n'est plus question d'analyser les phénomènes seulement à l'aide de « notions passives » (ou structuralistes) comme la diffusion et la réception, le centre et la périphérie, le national et l'international, le disciplinaire et l'interdisciplinaire mais il convient de s'intéresser aux circulations, aux réseaux, aux représentations, appropriations, assemblages, hybridations et territorialisations ; en bref, à ce qu'il y a de mouvant dans l'entreprise scientifique (Pestre, 1995 ; Anderson, 2002). Au bout du compte, le constructivisme des années 1970 et 1980 a incité à repenser les catégories d'analyse pour aborder les connaissances et les

pratiques scientifiques, autrement que par le modèle classique et plus statique qui nourrissait les raisonnements et les questions de recherche du cercle mertonien.

Pour prendre acte de ces évolutions, l'historien français des sciences physiques Dominique Pestre préconise l'utilisation du pluriel pour parler de science : « il est préférable d'abandonner la catégorie de Science, avec tout ce que celle-ci véhicule de réification, et de parler plutôt de champs disciplinaires et de pratiques matérielles et cognitives multiples. Les différences qui séparent la pratique de la philosophie naturelle de celles de la géologie du début du 19^e siècle et de la physique nucléaire de l'après Seconde Guerre mondiale sont telles que les subsumer sous une catégorie unique est un exercice trompeur conduisant des questions de peu de pertinence. Le postulat de l'unité des sciences appuyée sur une manière particulière de traiter les problèmes (ou son usage implicite à travers l'emploi de l'adjectif "scientifique") est certes revendiqué par de nombreux scientifiques mais rien n'oblige à le prendre pour argent comptant. Dit positivement, il semble préférable de considérer les divers moments et espaces comme autant de lieux aux règles épistémologiques, sociales, matérielles et rhétoriques spécifiques – chacun d'eux étant d'ailleurs plus une somme de systèmes et de représentations partiellement articulées qu'un ensemble unifié. Le rôle de l'historien est alors de suivre ces transformations, de construire une cartographie historique à plusieurs dimensions – et de ne point s'inquiéter outre mesure de la "coupure épistémologique". Historiciser radicalement la notion de science – comme écrire l'histoire de l'émergence rhétorique de cette notion – sont donc des préalables de ce programme. » (*ibid.*). À l'issue de ce bilan instructif quant aux apports du constructivisme pour les études sociales des sciences, Pestre envisage le développement de deux voies en histoire des sciences : la première concerne l'approfondissement de l'historiographie et la seconde (qui nous intéresse ici au premier chef) l'élaboration d'une géographie historique des sciences.

Voici la description qu'il fait de cette seconde voie : « L'une des "tendances de fond" les plus prometteuses actuellement est une approche qui spatialise les savoirs et construit une sorte de cartographie de géographie historique des pratiques savantes et techniques – et qui analyse la communication entre ces mondes hétérogènes *via* la standardisation des pratiques. Elle part de la multiplicité et de la diversité irréductible des lieux de production, elle admet l'hétérogénéité des régimes de légitimation, leur variabilité selon les espaces physiques et sociaux mais montre les moyens de normalisation mis en place par les acteurs pour échanger et progresser. » (*ibid.*). L'approche spatiale des sciences évoquée ici est apparue vers la fin des années 1980, et n'a fait que gagner des adeptes depuis, allant jusqu'à intéresser la géographie historique et culturelle (Naylor, 2005 ; R. C. Powell, 2007 ; Besse, 2010). À notre connaissance, l'un des premiers chercheurs à avoir attiré l'attention sur l'intérêt de cette approche spatiale des connaissances et

des pratiques scientifiques est le sociologue et historien des sciences Steven Shapin (Ophir & Shapin, 1991 ; Shapin, 1995, 1998). Issu de l'école d'Édimbourg (qui fut à l'origine du « programme fort »), il joue un rôle fondamental pour « redonner les pieds sur terre » à l'histoire des sciences qui, jusqu'au milieu du XX^e siècle, et avant le développement de la sociologie des sciences, n'était autre qu'une histoire décontextualisée des idées. Il participe dans cette perspective à réviser l'historiographie de la « révolution scientifique » (Shapin, 1996). Dès le début des années 1990, il se montre sensible à la tension entre l'ancrage spatio-temporel des savoirs (leurs productions et leurs appropriations sont situées) et le caractère « universel » que doivent présenter les savoirs « scientifiques » (Ophir & Shapin, *ibid.*). En effet, dans l'idéal universaliste, ce qui doit garantir à un savoir sa scientificité c'est son objectivité, c'est-à-dire sa capacité à être valide n'importe où et n'importe quand (Avramides, 2006). Cet idéal explique l'image, déjà évoquée, de « la tour d'ivoire » dans laquelle doivent exercer les savants : le laboratoire est généralement conçu comme un lieu détaché de son contexte environnant et des influences de ce dernier (Livingstone, 2003). C'est ce qui justifie que, pour comprendre comment des savoirs produits localement accèdent au statut de savoirs universels, les sociologues n'ont pas immédiatement privilégié une entrée spatiale.

En effet, les sociologues soucieux d'ouvrir la boîte noire des pratiques scientifiques n'ont pas véritablement prêté attention à la localisation des laboratoires qu'ils étudiaient (Latour & Woolgar, 1979 ; Latour & Callon, 1982). En particulier, l'espace n'est pas central dans les travaux de Bruno Latour, y compris quand il entend saisir le processus par lequel les chercheurs fabriquent des faits scientifiques, autrement dit, le processus permettant de transformer une observation localement située en énoncé scientifique universel (Latour, 1996). Néanmoins, les concepts développés par Latour (comme « l'acteur-réseau » et « la traduction »), ou encore l'analyse de John Law sur les méthodes des Portugais localisés à Lisbonne au XVI^e siècle pour contrôler et commander des territoires à distance (Law, 1986), posent des questions d'ordre géographique (Harris, 1998). En effet, ces approches invitent non seulement à ramener la production des énoncés scientifiques à leur contexte d'énonciation en tenant compte de la place des « non-humains » (les instruments par exemple), mais aussi à interroger les modes de circulation et de transmission des énoncés et des « objets intermédiaires » dans l'espace et le temps (cette dernière expression est empruntée à Dominique Vinck, 1999). Comme le résume Warwick Anderson, en analysant les effets du constructivisme sur l'histoire post-coloniale des technosciences, la position de Latour a même progressé à cet égard : « La théorie de l'acteur-réseau devait à l'origine fournir une explication regardant la production de ces "mobiles immuables", émergeant ainsi, presque paradoxalement, comme une variable non intentionnelle du vieux diffusionnisme ; les versions suivantes de la théorie ont mis en évidence

une topologie plus fluide, décrivant l'adaptation et la reconfiguration des objets et des pratiques en même temps qu'ils voyagent. »¹⁸ (Anderson, *op. cit.*).

Si c'est bien en faisant redescendre les idées et les savoirs « sur terre » que l'on peut s'occuper de la dimension spatiale de leur émergence et de leur circulation, ce n'est pas cette dernière qui intéresse Bruno Latour et ses confrères. Ces derniers cherchent à saisir les mécanismes de constructions des savoirs, non pas pour faire une géographie des savoirs mais pour étudier les échanges entre sciences et société. En fait, sans avoir jamais disparu, l'intérêt pour les rapports entre sciences et société se ravive au cours des années 1980, ce qui n'est pas étranger au contexte sociopolitique de cette décennie (abordé dans le Chapitre 2). Ce retour suppose de revaloriser les idées sinon du groupe des humanistes anglais de l'entre-deux-guerres, au moins du groupe allemand « Alternatives pour la science » fondé au sein du laboratoire du Max Planck Institut « Étude des conditions de vie dans un monde scientifico-technique ». Dans ce laboratoire créé à Starnberg en 1970 par Carl Friedrich von Weizsäcker et Jürgen Habermas, la science est envisagée comme une « ressource sociale », et l'on se soucie particulièrement de son utilité (Grossetti, 2008, p. 37). D'après Dominique Pestre, le groupe de Bruno Latour et Michel Callon participe à cette évolution en « déplaçant ses centres d'intérêt hors des milieux scientifiques ». En effet, pour comprendre « la science en action », l'approche de Latour évolue vers « un refus d'admettre qu'on puisse se limiter à l'univers des savants pour comprendre les sciences et leur dynamique, qu'on puisse se contenter des microanalyses de controverses », et Pestre de préciser : « non satisfait par la prétention sociologique d'expliquer un savoir (scientifique) par un contexte (social), Latour "sort" du laboratoire et cherche à comprendre comment le complexe (techno-)scientifique et le corps social se (re)définissent et se (re)construisent simultanément. Ce n'est plus localement, dans le seul cadre des laboratoires, qu'est cherché le secret des savoirs et de leur validation mais dans les reprises et traductions qui opèrent dans l'ensemble du corps social. » (Pestre, *op. cit.*). Sortir du laboratoire se légitime par la nécessité de saisir les relations des sciences avec la société, mais plus précisément avec la technologie et l'innovation, concepts qui prennent de plus en plus d'ampleur au moment de la crise des économies nationales dans les années 1970-1980. Dans ce contexte, Latour contribue à abattre les frontières épistémologiques, par exemple en forgeant la notion de « forums hybrides ». Cependant, il ne se saisit jamais d'entités spatiales pour traiter de la science en train de se faire, ce qui veut dire que son analyse du social n'est pas spatialisée.

¹⁸ L'original : « Actor-network theory initially was meant to provide an explanation for the production of these 'immutable mobiles', thus emerging, almost paradoxically, as an unintended variant of an older diffusionism; later versions have emphasized a more fluid topology, describing the adaptation and reconfiguration of objects and practices as they travel. » (Anderson, 2002).

En conséquence, même s'il en réunit toutes les conditions, ce n'est pas Latour qu'il faut convoquer pour comprendre l'entrée de l'espace, des lieux et territoires dans les études sociales sur les sciences. Cela dit, le parti pris latourien se répercute sur l'approche spatiale défendue par Shapin pour étudier : la fabrication, la transformation et la circulation des savoirs, des techniques et des objets scientifiques d'une part ; et/ou pour explorer les liens entre mondes scientifiques, économiques et politiques d'autre part. Plusieurs voies ont été ouvertes pour y parvenir, notamment en France.

L'historien Christian Jacob est à l'origine de l'une de ces voies : « l'anthropologie historique et comparée des pratiques savantes » (comme l'indique le sous-titre de son carnet de recherche en ligne *Lieux de Savoir*). Depuis les années 1980, l'ensemble de son œuvre fleuve est habitée d'une réflexion sur les liens entre espace et savoirs. Sa trajectoire au sein de l'École des Hautes Études en Sciences Sociales (EHESS) l'a amené à être influencé par Jean-Pierre Vernant, son mentor en anthropologie historique de l'Antiquité, par l'épistémologue Emile Meyerson (maître de ce dernier), par le philosophe Michel de Certeau, mais aussi par les historiens américains de la cartographie Bryan Harley et David Woodward¹⁹. Comme un certain nombre de promoteurs d'une géographie des connaissances (David N. Livingstone, Charles W.J. Withers, Richard C. Powell, Simon Naylor, Steven J. Harris...), c'est en partant de l'histoire de la géographie que Jacob en est venu à s'intéresser à la géographie des savoirs. Plus précisément, depuis sa thèse d'État saluée par Paul Claval (Claval, 1988), son approche théorique de la cartographie à travers l'histoire l'a conduit à interroger la place des lieux pour les pratiques savantes. S'il est moins ambitieux, le travail du sociologue des connaissances australien David Turnbull, proche de l'historien des sciences David W. Chambers, procède du même cheminement (Turnbull, 1996, 1997). En dépit de cette ressemblance, les figures de référence menant à la géographie des sciences varient d'un auteur à l'autre. En particulier, alors qu'à l'étranger, le rôle des géographes post-modernes (Derek Gregory, David Harvey, Denis Cosgrove ou encore Allan Pred) est jugé essentiel pour les tenants d'une approche spatialisée des savoirs (Withers, 2009), ce sont plutôt les épistémologues et philosophes des sciences qui sont désignés en France comme étant au fondement de cette démarche (par exemple Michel Serres cité par Jean-Marc Besse, *op. cit.*). Il faut dire que le post-modernisme ou « tournant culturel » a constitué un moment important de la géographie anglo-américaine, notamment en légitimant son intérêt pour l'imagination, les idées et les représentations, ce qui *a poste-*

¹⁹ Cette liste de personnalités a été établie à partir de la réponse apportée par Christian Jacob à l'une de nos questions lors de la séance de clôture du séminaire de l'équipe PLH (Patrimoine, Littérature, Histoire) E.R.A.S.M.E (Équipe de recherche sur la Réception de l'Antiquité : Sources, Mémoires, Enjeux) qui s'est tenue le 16 avril 2014 à la Maison de la Recherche de l'Université de Toulouse, campus du Mirail.

riori apparaît décisif pour le développement d'une géographie des connaissances scientifiques dans le champ anglophone (Chapitre 1.2).

Ainsi, lorsqu'au cours des années 2000, Christian Jacob tâche de fédérer un ensemble de chercheurs français autour du thème des lieux de savoirs, ce sont surtout des spécialistes des sciences, historiens pour la plupart qui se joignent à son entreprise intellectuelle. Celle-ci débouche sur le premier ouvrage collectif de la série *Lieux de savoirs*, sous-titré *Espaces et Communautés* (Jacob, 2007). Dans la recension qui en est faite par Nicolas Adell-Gombert, deux entrées sont distinguées : l'entrée par les savoirs dont on étudie la mise en espace d'une part, et l'entrée par le lieu où les savoirs sont situés d'autre part (Adell-Gombert, 2008). À l'honneur dans la dernière partie du livre « Ville Phare », l'entrée par les lieux établit un pont entre les études urbaines et les études sociales des sciences. Elle intègre le travail des historiens Antonella Romano et Stéphane Van Damme qui, à la suite d'un programme de recherche ANR Jeunes Chercheurs, ont coordonné le numéro spécial « Sciences et Villes-Mondes, XVI^e – XVIII^e siècles » (Romano & Van Damme, 2008). En attirant l'attention sur « l'espace public » et en plaçant le niveau d'analyse du travail scientifique à l'échelle urbaine, ce projet envisage la ville – au même titre que le laboratoire, l'université, l'observatoire ou le jardin botanique – comme un lieu de science. Ce faisant, il invite à considérer le niveau urbain comme un niveau pertinent pour étudier les pratiques scientifiques localisées. Dans le même registre mais élargi à l'ensemble des activités culturelles bénéficiant d'un fort rayonnement, on peut penser au projet collectif dirigé par Christophe Charles autour des capitales culturelles (Charles, 2009). Aussi, puisqu'ils sont tournés vers le passé, ces projets rappellent le projet de sociohistoire des pôles scientifiques amorcé par Michel Grossetti et ses confrères dans les années 1990 (déjà abordé en Chapitre 3.2).

Avec une approche plus sociologique qu'anthropologique, Grossetti *et al.* envisagent très tôt le niveau urbain comme idéal pour saisir la question des interrelations entre sciences et sociétés et cela jusqu'à l'époque contemporaine. Preuve empirique à l'appui, Michel Grossetti s'achemine vers cette approche en commençant par montrer, au cours des années 1990, dans quelle mesure le niveau local suivi du niveau national sont structurants pour l'établissement de relations science-industrie à l'échelle française (Grossetti, 1995). Il faut dire que cette question était particulièrement d'actualité au moment où faisait irruption sur la scène académique l'ouvrage dirigé par Michael Gibbons sur le nouveau mode de production du savoir des sociétés contemporaines (le « mode 2 ») (Gibbons *et al.*, 1994). Cet ouvrage a poussé à leur paroxysme les théories « anti-différenciationnistes » (l'expression vient de Shinn & Ragouet, 2005) considérant que les frontières entre sciences et société, entre disciplines et entre États n'avaient plus de sens dans le cadre d'un régime économique mondialisé. Puisque l'image du monde qu'il véhi-

culait tenait du manifeste politique, cet *opus* n'a pas convaincu les spécialistes des sciences. Dès sa sortie, de nombreux détracteurs l'ont employé pour en explorer les limites sur leurs terrains d'étude. Loet Leydesdorff et Henry Etzkowitz ont proposé, sur cette base, un modèle plus souple et réaliste de la place de la science en société : le modèle de la triple hélice. Inspiré de la structure du brin d'ADN, ce dernier modèle invitait à considérer simultanément les relations entre science, industrie et politique. En 2002, Terry Shinn (autrefois membre du groupe Pandore, mais institutionnaliste), propose une critique des deux modèles (celui du nouveau mode de production du savoir et celui de la triple hélice) (Shinn, 2002). Même si leurs prétentions diffèrent (le second a plus une visée scientifique et opératoire que le premier), ces deux modèles apparaissent beaucoup trop simplistes, alors que d'après Shinn, « on doit s'assurer au maximum que les concepts et les conclusions s'enracinent solidement dans des études empiriques, que les conditions limites des résultats sont bien établies et que les modèles reflètent de façon prudente et critique les contraintes des données. » (*ibid.*).

L'absence de fondement empirique est particulièrement alarmante dans le manifeste de Gibbons *et al.* : « les assertions avancées ne reposent presque jamais sur des informations concrètes » (*ibid.*). Dans les années 1990, cette critique s'applique à de nombreux travaux qui entendent saisir le nouvel ordre mondial. De manière analogue et au sujet de l'ouvrage de Roland Robertson sur la globalisation culturelle, Yves Lambert déclare : « Malgré ces précieux apports, l'ouvrage présente les inconvénients de beaucoup d'écrits sur la globalisation : discours trop général, usage excessif de notions vagues, redondance des analyses (...) et surtout, absence d'enquête sociologiques. » (Lambert, 1995). En contrepoint, l'idée que c'est aux sciences sociales de saisir, à l'aide de données empiriques adaptées, les grandes tendances mondiales en s'affranchissant des contraintes, déjà évoquées, dont souffrent les statistiques publiques se répand progressivement. Pour mieux saisir cette réalité en mouvement, « l'approche réseau » s'impose comme une alternative aux approches structuralistes (centre-périphérie) désormais jugées trop statiques. Très précisément, Jacques Gaillard et Jean-Baptiste Meyer en font la démonstration pour l'étude des migrations scientifiques (Gaillard & Meyer, 1996). Au lieu d'estimer les flux de chercheurs entrants et sortants des pays, ils montrent la nécessité d'identifier les relations « transnationales » qui se mettent en place au sein de diasporas, autrement dit, les relations entre chercheurs partageant la même nationalité mais répartis sur plusieurs pays à la fois. D'ailleurs, c'est dans ce même mouvement que la notion de « communautés » réapparaît, mais pour désigner des collectifs plus souples qu'auparavant.

Finalement, deux phénomènes caractérisent le rapport au temps et à l'espace des travaux sur la science à partir des années 1990. Le premier est la volonté de répondre à la question qui se trouve, d'après Laurent Rollet, au cœur des démarches, pourtant bien

distinctes, de Christian Jacob et Michel Grossetti : « Dans une certaine mesure, il serait possible de considérer que l'étude historique des pôles scientifiques se résume à une question générale et non spécifique à cette thématique : comment des savoirs en viennent-ils à faire corps, à faire lieu, à être partagés dans des collectifs, à organiser des territoires, à circuler dans des réseaux ? » (Rollet, 2009). Le second est la nécessité « d'inventer les bases de nouvelles données » (pour reprendre encore Wallerstein, 1982) afin de saisir plus finement et en dynamique l'état de la science dans le monde. Dans cette optique, une méthode est amenée à connaître d'amples développements parce qu'elle permet de confronter à la réalité les représentations toutes faites sur l'inscription spatiale des pratiques scientifiques : la « scientométrie spatiale ».

2.2. ...à la « scientométrie spatiale »

Comme nous l'avons déjà évoqué, un moyen courant de mesurer et de localiser l'activité scientifique depuis le début du XX^e siècle consiste à se pencher sur les produits de cette activité. Parmi ces produits, les publications fournissent des informations intéressantes d'un point de vue géographique puisque le nom des auteurs s'accompagne de leur adresse institutionnelle. La démarche consistant à considérer l'information spatiale contenue dans les données bibliographiques a fait l'objet d'un nombre croissant de travaux au cours des années 2000. Cela tient au fait que les possibilités de traitement automatiques se sont améliorées mais aussi que l'intérêt pour la géographie des activités scientifiques et innovantes a progressé.

Le traitement spatialisé d'ensembles de données bibliométriques requiert une sensibilité aux biais caractéristiques des sources bibliographiques. Parce qu'il exige la combinaison de savoir-faire en scientométrie et en analyse spatiale agrémenté d'une familiarité avec la littérature portant sur les sciences et l'innovation, cet enjeu en est venu à attirer l'attention d'un spécialiste néerlandais de l'innovation localisé à Utrecht, Koen Frenken. Il préconise en 2009, avec ses doctorants Sjoerd Hardeman et Jarno Hoeckman, le développement d'un nouveau programme de recherche baptisé « scientométrie spatiale » (Frenken, Hardeman, & Hoekman, 2009). Ce programme se justifie par la nécessité de tirer parti des progrès réalisés en traitements de l'information et en visualisation assistés par ordinateur. Aussi, peu de temps après cette parution, Van Noorden s'en inspire pour un article de vulgarisation publié dans *Nature* qui met l'accent sur le thème « villes et sciences » (Van Noorden, 2010). Comme les travaux se revendiquant de ce courant ont été nombreux depuis la parution de l'article de Frenken et ses co-auteurs, ces derniers se sont occupés de recenser les évolutions intervenues entre temps et les perspectives à venir (Frenken & Hoekman, 2014).

Évidemment, les contributions des équipes spécialisées en scientométrie ou sciences de l'information jouent un rôle important sur ce front. Déjà, jusqu'au milieu des années 2000, les scientomètres avaient pour habitude de réaliser des traitements géographiques au niveau national. Ce niveau était privilégié à la fois parce que c'est celui auquel se font les politiques mais aussi parce que traiter l'information à ce niveau présente moins de difficultés méthodologiques – si bien qu'il existe un très riche ensemble de travaux, pour une grande partie d'entre eux publiés dans la revue *Scientometrics* à partir de la fin des années 1970, qui mesurent la production et l'impact scientifique par pays ainsi que l'évolution des collaborations internationales. Cet ensemble de littérature sera étudié plus en détail lorsque nous aborderons l'histoire et le contenu de la source bibliographique privilégiée par les scientomètres : le *Web of Science*.

Dans ce cadre, les scientomètres néerlandais sont à la pointe pour essayer, à l'aide de nouveaux moyens techniques, de traiter, d'analyser et de visualiser plus finement et plus rapidement l'information géographique contenue dans les notices bibliographiques. À Amsterdam, Loet Leydesdorff s'intéresse particulièrement à ce problème. En 2010, il co-publie un article avec le scientomètre suédois Olle Persson dans lequel il aborde le thème de la géolocalisation automatique des adresses de publications, la représentation automatique des réseaux scientifiques de villes et d'institutions (Leydesdorff & Persson, 2010). À partir de 2010, il entame une collaboration avec le scientomètre allemand Lutz Bornmann sur les enjeux méthodologiques associés à la « cartographie de l'excellence » (Bornmann, Leydesdorff, Walch-Solimena, & Ettl, 2011 ; Bornmann & Leydesdorff, 2012). Bornmann collabore également avec Ludo Waltman au centre d'études de la science et technologie à Leyde, qui se spécialise dans les questions de visualisation et de traitement de l'information géographique des données bibliographiques (Waltman, van Eck, & Noyons, 2010 ; Bornmann & Waltman, 2011). Sur ces questions de visualisation ou *science mapping* explorées dans la seconde partie de la thèse, on peut aussi citer les contributions du groupe de recherche espagnol SCImago, dirigé par Felix de Moya Anegón, associé à l'Université de Grenade et au CSIC (Olmeda-Gomez, Perianes-Rodriguez, Antonia Ovalle-Perandones, Guerrero-Bote, & de Moya Anegón, 2009 ; Chinchilla-Rodriguez, Vargas-Quesada, Hassan-Montero, Gonzalez-Molina, & Moya-Anegón, 2010) et les contributions menées sous la houlette de Katy Börner du groupe américain en Science de l'information de l'Université de Bloomington en Indiana (Boyack, Klavans, & Börner, 2005 ; Börner, Penumarthy, Meiss, & Ke, 2006).

Prochainement, on s'attend à voir d'autres laboratoires de scientométrie et sciences de l'information participer ou tirer parti de ces avancées. En particulier, les structures à cheval sur la recherche et la production de statistique publique telles que les Observatoires des Sciences et Techniques (OST) en France et à Montréal sont, on

l'expliquera dans le prochain chapitre, au fait des efforts qui ont été réalisés dans ce domaine par notre équipe. Enfin, caractéristiques des tendances récentes en matière de traitement et de visualisation de l'information bibliographique, le rapport réalisé sous la direction de Caroline Wagner pour la Royal Society en 2011 montre l'évolution de la production scientifique dans les grandes villes mondiales et s'accompagne d'une galerie de graphes de collaborations internationales mais aussi d'une carte (Figure 9) (Royal Society, 2011).

Les 20 premières villes en 2004-2008, et leur croissance depuis 1996-2000



Clef Les villes responsables du plus grand nombre de publications en 2004-2008; croissance depuis 1996-2000

- Décroissante ou constante
- A gagné 5 à 10 places
- A gagné 10 à 20 places
- A gagné plus de 20 places

Figure 9 – Carte conçue par Elsevier à partir des données de la base Scopus. Extraite et traduite du rapport « *Knowledge, networks and nations* » (Royal Society 2011, p. 38)

Outre les contributions des scientomètres à la « scientométrie spatiale », considérons enfin les expériences de chercheurs intéressés par les rapports entre science et espace sans que leur spécialité dominante soit la scientométrie. Dans cet ensemble, on distingue trois sous-ensembles qui ne sont pas vraiment isolés les uns des autres, que nous avons

tour à tour abordés au cours de l'état de la littérature : l'économie régionale et la géographie de l'innovation, les études sociales des sciences, l'histoire et la géographie des sciences.

En économie régionale et géographie de l'innovation, les travaux mobilisant des données bibliométriques sur la science académique sont encore rares. À Toulouse, c'est à la suite des études spatialisées des contrats CNRS-entreprise (Grossetti & Nguyen, 2001) que Béatrice Milard a eu l'idée de mesurer pour la première fois la traduction spatiale de l'hybridation entre organismes nationaux de recherche et universités à partir des données de publications (Grossetti & Milard, 2003). Cette initiative s'inscrit dans une dynamique plus générale et on peut à nouveau rappeler les thèses récentes réalisées à Poitiers autour d'Olivier Bouba-Olga par Marie Ferru (Ferru, 2009) et Bastien Bernala ; et aux Pays-Bas, les travaux engagés autour de Koen Frenken et Ron Boschma par Roderik Ponds (Ponds, 2008) et Pierre-Alexandre Balland (Balland, 2011). Le cadre théorique de référence pour ces derniers travaux est celui de l'école de la proximité, autrement dit, la mesure des effets des différents types de proximité (géographique, organisationnelle, cognitive...) sur les échanges scientifiques observés. De plus en plus souvent, cette approche s'accompagne d'un recours aux outils de l'analyse de réseaux, en témoigne la thèse de David Frachisse dirigée par Nadine Massard (Frachisse, 2011) et celle de Mickaël Benaim par Jean-Alain Héraud (Benaim, 2013).

Dans le second sous-ensemble, celui des études sociales sur les sciences, on peut signaler les études spatialisées des logiques d'organisation scientifiques, y compris au sein de sous-domaines disciplinaires ou spécialités scientifiques, comme celles que Leydesdorff a réalisées en collaboration avec la spécialiste américaine des politiques publiques de la science Caroline Wagner (Wagner, 2005 ; Wagner & Leydesdorff, 2005) ou encore avec Ismael Rafols, spécialiste de l'interdisciplinarité et des effets d'émergence (Leydesdorff & Rafols, 2011). Aussi, il importe de considérer les études spatialisées de trajectoires ou carrières scientifiques, et notamment celles que Grit Laudel a réalisées sur la mobilité géographique de chercheurs « stars » (Laudel, 2005). En effet, cette sociologue allemande est l'une des premières à avoir analysé la mobilité géographique de chercheurs à l'aide de traitements bibliométriques (Laudel, 2003). Elle s'est aussi consacrée à une réflexion sur l'intérêt pour la sociologie des sciences de renouer avec la scientométrie, après s'en être éloignée au cours des années 1980-1990 (Gläser & Laudel, 2001).

Pour finir, on trouve de plus en plus de géographes et d'historiens intéressés par la « scientométrie spatiale ».

En géographie, il y a les études déjà évoquées sur les villes mondiales de la science (Matthiessen & Schwarz, 1999 ; Matthiessen, Schwarz, & Soren, 2002 ; Matthiessen, Schwarz, & Find, 2002 ; Matthiessen *et al.*, 2010), les systèmes de villes européens (Co-

min, 2009), l'histoire de la géographie (Cuyala, 2014) et la géohistoire de la science moderne (Taylor *et al.*, 2008). Ces contributions témoignent de l'inclination des géographes à prendre la ville comme unité d'analyse pour étudier les flux scientifiques. Par ailleurs, il importe de signaler les contributions de Stefan Henneman et ses co-auteurs (Hennemann, Rybski, & Liefner, 2012 ; Hennemann, 2013). Celles-ci montrent tout l'intérêt d'une maîtrise en matière de représentation cartographique des données et d'un savoir géographique pour aborder le lien entre science et espace (2013). Puisqu'il interroge le « mythe de la science globale » (2012), Henneman participe d'un ensemble d'études engagées depuis le début des années 1990 sur la mondialisation de la science à l'aide de données bibliométriques (Schott, 1993). Récemment, dans cette catégorie à laquelle se rattache la présente thèse, s'inscrivent la thèse de Hoekman, soutenue en 2012, qui traite de la globalisation à travers la question des essais cliniques (Hoekman, 2012) ; mais aussi la thèse du géographe russe Koval'ev qui est le signe, s'il en fallait encore, qu'un mouvement généralisé à de nombreux pays, depuis le début des années 2000, conduit les géographes à s'intéresser aux activités scientifiques (Kovalev, 2002).

En histoire des sciences, moins courante que l'approche qualitative, l'approche quantitative tend à se développer en même temps que les possibilités de traitement numérique des sources. Le travail monumental de l'historien des sciences anglais Robert Mortimer Gascoigne (1918-1994) qui s'est attelé à constituer des catalogues de revues et de scientifiques européens à partir du XVII^e siècle et une chronologie d'histoire des sciences qui remonte au XV^e siècle – est très précieux dans cette perspective (Gascoigne, 1992). Peter Taylor mobilise cette source pour son analyse géohistorique de la science moderne (*op. cit.*) et, dans le cadre du programme « Géoscience », l'historien des sciences René Sigrist l'a exploitée pour étudier la géographie de plusieurs « communautés savantes » de chimistes à partir du XVII^e siècle (Sigrist, 2013). La volonté de traiter systématiquement ou quantitativement les marqueurs ou traces de l'activité scientifique issues des sources écrites se retrouve chez les historiens des sciences, y compris quand ils ne sont pas modernistes. Par exemple, l'historien de l'Antiquité Anthony Andurand étudie les lieux et les connexions repérables dans *Les propos de table* de Plutarque afin de mieux saisir l'organisation du monde savant gréco-romain (Andurand, Jégou, Maisonobe, & Sigrist, 2014).

La division du travail à laquelle nous venons de procéder est insatisfaisante parce que « la scientométrie spatiale » est un programme transversal, de l'ordre de la proposition méthodologique. Ici, en reprenant des travaux déjà cités ou en citant des travaux qui ne l'avaient pas encore été, nous avons montré que cette démarche attire à la fois des spécialistes du traitement d'information et de la visualisation que des chercheurs en sciences sociales qui s'intéressent au lien entre sciences et espace pour affronter des problèmes

d'économie, de sociologie, de géographie et d'histoire. Or, que ce soit pour leur dimension méthodologique ou pour leur dimension théorique, ces travaux sont en mesure d'intéresser les géographes et en mobilisent déjà quelques-uns.

Plus généralement, ce détour par les *science studies* montre combien la géographie gagnerait à s'atteler à la construction d'un cadre théorique unifié autour de la question des sciences. La géographie des sciences qui reste à définir devrait être capable de concilier sous une même bannière les travaux qualitatifs à dimension historique, réalisés sous l'influence de Steven Shapin, et les études quantitatives relevant de la « scientométrie spatiale ». Dans cette perspective, le programme de recherche « Géoscience » – abordé dans le prochain chapitre – a été éclairant. Les forces qu'il a rassemblées, issues de plusieurs disciplines, ont permis de creuser dans plusieurs directions enrichissant l'ensemble des problématiques et des enjeux méthodologiques qu'une géographie des sciences serait en mesure de couvrir. À l'issue de ce programme, l'opération « mondes scientifiques » du laboratoire d'excellence « Structuration des Mondes Sociaux » (Labex SMS) entamée en 2013 est une invitation à poursuivre les réflexions en allant vers une théorie plus vaste encore²⁰. Il s'agit effectivement d'envisager les articulations et les recoupements entre les différents « mondes sociaux » (au sens d'Howard Becker). Partisan de cette formulation moins contraignante que « communauté », Michel Grossetti, qui compte parmi les artisans du projet SMS, a convoqué dès le début des années 2000 avec Marie-Pierre Bès les figures de Karl Polanyi, Harrison White et Marc Granovetter pour attirer notre attention sur les phénomènes d'« encastrement » et de « découplage » qui résultent des différences entre les « mondes », et continue de s'y référer pour comprendre les phénomènes d'émergence de collectifs sociaux notamment dans les mondes des arts et des sciences (Bès & Grossetti, 2003 ; Grossetti & Bès, 2002 ; Grossetti, 2007, 2015). Pour imager, il souligne par ce chantier l'intérêt d'appréhender en dynamique les successions d'harmonies et de dissonances (les bifurcations) dans les partitions jouées simultanément par les individus, collectifs et institutions en sociétés.

²⁰ Voir le texte décrivant le projet du Labex SMS : LABEX SMS. (2013). *Présentation du Labex SMS* (Projet de recherche). Toulouse : Université de Toulouse. 23 p. Consulté à l'adresse <http://sms.univ-tlse2.fr/accueil-sms/le-labex-sms/>

CHAPITRE 5. CADRE ET SUPPORT DU TRAVAIL : LE PROGRAMME « GÉOSCIENCE » ET LE WEB OF SCIENCE

Au début des années 1990, le sociologue américain Thomas Schott contribue à connecter le thème de la « globalisation » à celui des sciences dans un article intitulé : « *World Science : Globalization of Institutions and Participation* » (Schott, 1993). À la fin de l'article, dans la section « Étudier la science mondiale », Schott suggère que le phénomène est mesurable, nous incitant à adopter une approche de scientométrie spatiale pour l'étudier : « La participation d'une communauté scientifique peut être mesurée à l'aide de nombreux concepts, dont la production de connaissance, l'influence intellectuelle et la collaboration. Par chance, des indicateurs pour ces trois variables peuvent être construits à partir d'une source d'archive accessible, la littérature scientifique, qui a l'avantage de couvrir le système global. Les publications indiquent la production de connaissance dans une communauté, les citations indiquent l'influence intellectuelle, et les co-publications indiquent les collaborations. »²¹. 20 ans plus tard, les chercheurs se rassemblant autour du programme « Géoscience » envisagent l'exploitation de ce type de sources pour saisir à la fois les logiques spatiales qui participent à faire exister et grandir des « communautés scientifiques nationales », mais aussi les logiques de participation d'un nombre croissant de scientifiques à l'ensemble de la « science mondiale ». Parallèlement, pour enrichir leur compréhension de la géographie des activités scientifiques à plusieurs niveaux spatiaux, les membres du programme intègrent la nécessité de se nourrir d'observations, d'entretiens et de recherches plus qualitatives, tout en privilégiant un éclairage historique.

Définis dans le cadre de ce programme de recherche, les enjeux de la présente thèse ne peuvent être compris qu'au regard du travail d'équipe pluridisciplinaire et des réflexions auxquelles il a donné lieu. C'est pourquoi, en revenant dans ce chapitre sur le contenu et le déroulement du programme de recherche « Géoscience », nous allons nous

²¹ L'original : « Participation of a scientific community can be measured in terms of several concepts, including knowledge production, intellectual influence, and collaboration. Fortunately, indicators for these three variables can be constructed from an accessible archival source, the scientific literature, which has the advantage of covering the global system. Publications indicate knowledge production in a community, citations indicate intellectual influence, and coauthorship indicates collaboration. » (Schott, 1993).

efforcer de mettre en évidence les liens existants entre notre projet doctoral et les avancées et propositions, notamment méthodologiques, élaborées collectivement.

Après avoir posé le décor et avant d'entrer dans la deuxième partie de la thèse, qui traite du protocole de scientométrie spatiale mis au point en équipe et des grandes tendances qu'il nous a permis de mettre à jour, il conviendra de se pencher sur les origines et le contenu de la source bibliographique de notoriété mondiale ayant servi à établir ce protocole : le *Web of Science* (WoS). Puisque notre équipe s'est appuyée sur cette source pour établir une méthode inédite et l'a privilégiée pour saisir les grandes tendances contemporaines de la géographie des activités scientifiques, il importe d'en connaître les avantages et les inconvénients. Une fois que les biais et les limites de la source seront connus, que sa provenance et sa fiabilité seront identifiées, il sera plus évident de comprendre ce qu'elle est en mesure de montrer, de déterminer comment s'y prendre pour l'exploiter, et d'interpréter les données géographiques qu'elle recèle.

1. Le programme « Géoscience »

« L'aventure commença par une mission des plus banales, la routine, le quotidien, l'ordinaire. Il y avait des années que le travail sur le continent antarctique n'était plus l'affaire des intrépides, mais celle des sages organisateurs. On avait tout le matériel qu'il fallait pour lutter contre les inconvénients du climat et de la distance, pour connaître ce qu'on cherchait à savoir, pour assurer aux chercheurs un confort qui eût mérité au moins trois étoiles – et tout le personnel nécessaire possédant toutes les connaissances indispensables. » (Barjavel, 1968, p. 11)

1.1. Une approche mixte et interdisciplinaire

À cette étape, le programme de recherche ANR « Géoscience » mérite d'être abordé en détail pour au moins trois raisons : premièrement, parce que notre implication dans ce programme est partie prenante de la thèse, deuxièmement parce qu'en rassemblant de nombreux chercheurs de différentes disciplines, « Géoscience » est un moment important des études sociales des sciences en France, et troisièmement parce que les apports de « Géoscience » du point de vue de la scientométrie spatiale sont inédits. En effet, dans le cadre du programme sous-titré « Science locale, nationale, mondiale en transformation. Pour une socio-géographie des activités et des institutions scientifiques acadé-

miques », le traitement de l'information spatiale contenue dans le WoS a été réalisé à un niveau de précision jamais égalé : le niveau urbain ou local. Pour l'exploitation à ce degré de précision de l'information à l'échelle mondiale, il a fallu mettre au point, en équipe, une méthodologie robuste. Puisque nous y avons contribué, cette progression méthodologique est tout à la fois un résultat de la thèse et le socle sur lequel reposent les observations empiriques qui y sont analysées à la fois sur l'ensemble de la production scientifique mondiale et dans le cadre d'une étude de cas réalisée de façon plus autonome.

Avant d'en venir à la présentation du protocole élaboré collectivement et à la restitution des résultats d'équipe (Partie 2), la présentation du programme de recherche est un arrêt obligé, y compris pour éclairer le parti pris théorique de cette thèse et expliciter ce que nous entendons par « géographie des sciences ». En effet, même si des progrès majeurs ont été réalisés en scientométrie spatiale, la partie purement quantitative ou statistique (la « tâche 2 ») n'était qu'une des dimensions du programme. Bien que ce soit la partie dans laquelle nous avons été le plus sérieusement impliquée, il ne faudrait pas laisser penser que la géographie des sciences devrait être abordée uniquement par le biais de la scientométrie spatiale. Comme en témoigne la fiche extraite d'un document retraçant les intentions du projet de recherche (Tableau 2) et le schéma récapitulatif (Figure 10), « Géoscience » s'est donné l'ambition d'articuler une analyse historique de la construction des espaces nationaux, avec une focalisation particulière sur le cas français (tâche 1), une étude quantitative de la répartition dans l'espace (aux niveaux national et international) des publications scientifiques et des collaborations (tâche 2) et une série d'études de cas de collaborations internationales (tâche 3).

L'esprit mixte dans lequel s'est déroulé le programme, autrement dit la volonté à tous les instants de faire co-exister les analyses quantitatives et qualitatives ; mais aussi l'intérêt des participants pour les analyses multi-niveaux – le niveau des villes ou des trajectoires individuelles, le niveau des régions ou des pays et le niveau mondial – sont autant de dispositions qui furent d'un apport considérable. De façon analogue – puisque l'espace scientifique a en commun avec l'espace géographique d'être un espace hétérogène, fragmenté et hiérarchisé – les circulations ont été fréquentes entre le niveau macro-scientifique (toutes disciplines confondues), le niveau méso-scientifique (les disciplines ou spécialités) et le niveau micro-scientifique (les trajectoires individuelles ou les questions de recherche).

GEOSCIENCE	Science locale, nationale, mondiale en transformation. Pour une socio-géographie des activités et des institutions académiques			
Début ►	janv-10	Fin ►	juin-13	Durée ► 42 mois
Structure de Financement	ANR, appel d'offre « Sciences et Société » (2009)			
Coordination générale	Michel Grossetti et Denis Eckert (UMR LISST, UMR EHESS et Université de Toulouse-le Mirail)			
Partenaires français	<p>1. Laboratoire Interdisciplinaire Solidarités, Solidarités, Territoires (LISST, UMR5193) Toulouse (Michel Grossetti et Denis Eckert);</p> <p>2. Centre Maurice Halbwachs (CMH, UMR 8097), Paris (André Grelon);</p> <p>3. IRD (UMR 201), équipe Science, Technologie et Société (STS), Paris, (Rigas Arvanitis, Mina Kleiche-Dray, Jacques Gaillard, Roland Waast); Luigi Rossi (IRD-DIC)</p> <p>4. Politiques Publiques, Action Publique, Territoires (PACTE, UMR 5194), Grenoble, (Dominique Vinck).</p>			
Partenaire étranger	5. Centre Interdisciplinaire de Recherche sur les Sciences et les Techniques, CIRST, UQAM, Montréal (Yves Gingras)			
Description du projet	<p>Le projet GEOSCIENCE a pour objectif d'éclairer les débats publics sur l'évolution des espaces d'exercice de l'activité scientifique à partir d'une articulation étroite entre les études sociales sur les sciences (sociologie des sciences, histoire des sciences et des techniques, économie de l'innovation) et la géographie urbaine. GEOSCIENCE vise à étudier les logiques spatiales actuelles de l'activité scientifique au niveau local (villes, régions), national et international, en les articulant à une analyse historique de la construction des institutions scientifiques et de leur déploiement dans l'espace durant les deux derniers siècles. Le projet GEOSCIENCE se centrera particulièrement sur la tension existant entre, d'une part, les logiques de différenciation et de hiérarchisation des espaces de production scientifique (affirmation des centres anciens ou de grande taille au niveau infra-national, des pays « dominants » au niveau international), et d'autre part les logiques de rééquilibrage (multiplication des centres atteignant une taille importante, émergence au niveau international de nouvelles « puissances scientifiques »). L'articulation de la perspective historique et de l'analyse contemporaine se fera à partir de la mobilisation d'analyses historiques, la réalisation d'études quantitatives approfondies, principalement sur les données bibliométriques, et la conduite d'études qualitatives de type sociologique. Le projet, d'une durée de 3 ans, articulera une analyse historique de la construction des espaces nationaux, avec une focalisation particulière sur le cas français (tâche 1), une étude quantitative de la répartition dans l'espace (aux niveaux national et international) des publications scientifiques et des collaborations (tâche 2) et une série d'études de cas de collaborations internationales (tâche 3). Il associera pour cela des sociologues des sciences, des géographes, des économistes et des historiens. En effet, ce projet se veut aussi l'occasion de construire une socio-géographie de l'activité scientifique qui tire parti des acquis des études sur l'histoire des institutions scientifiques, de la scientométrie, de l'analyse sociologique des collaborations scientifiques et de la géographie régionale et urbaine.</p>			

Tableau 2 – Fiche-résumé du projet « Géoscience ». Tableau repris d'un document de travail réalisé par Mina Kleiche-Dray

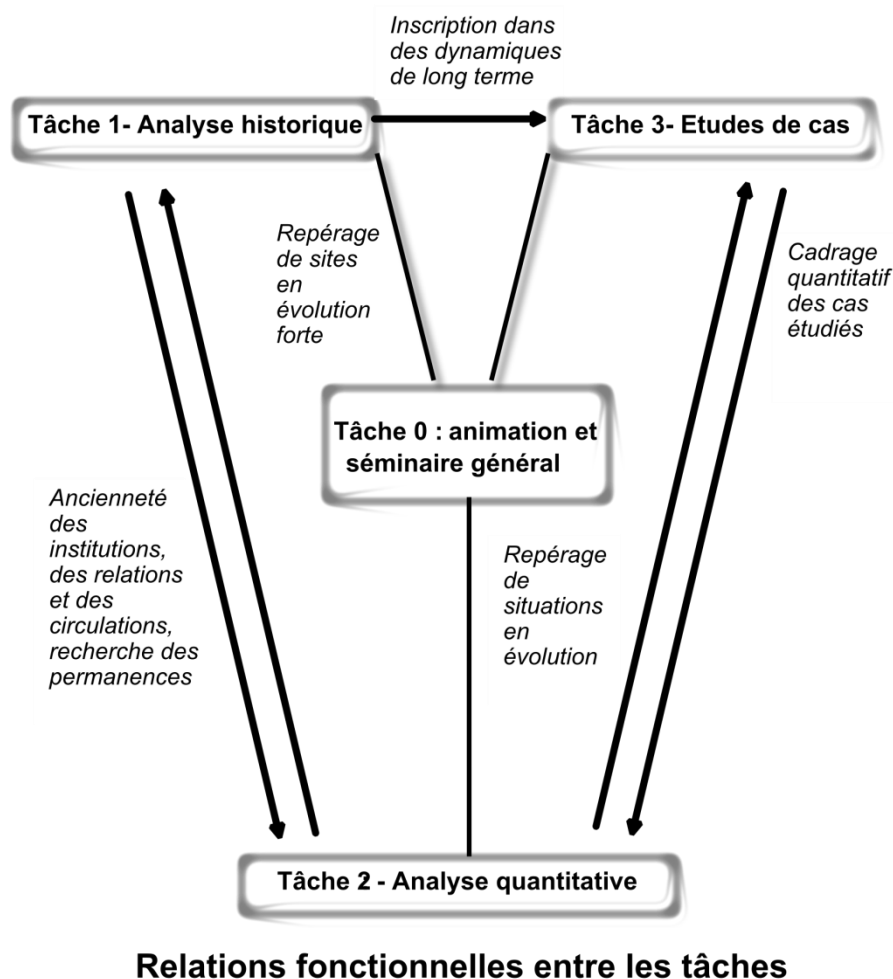


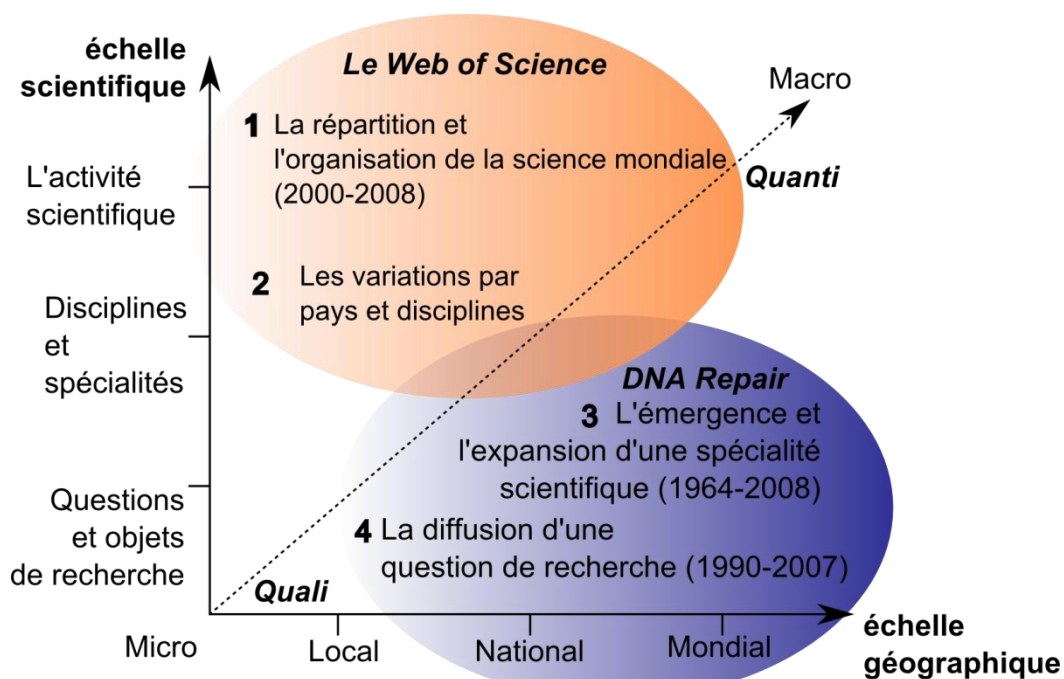
Figure 10 – Les relations fonctionnelles entre les différentes tâches du programme « Géoscience ». Schéma repris d'un document de travail

En parallèle, ces dispositions nous ont incitées à concevoir une thèse qui ne soit pas purement quantitative et axée sur les grandes tendances de l'activité scientifique au niveau mondial. Souhaitant également articuler différents niveaux d'observation, nous avons pris le parti de mettre en évidence et de caractériser des objets relevant de différents niveaux géographiques et scientifiques à la fois :

- les grandes tendances macro scientifiques (tous domaines confondus) et géographiques (mondiales) regardant la répartition des activités scientifiques et son organisation (Partie 2) ;
- les différentes manifestations de ces tendances selon les pays et les disciplines (Partie 2) ;

- le développement d'une spécialité en biologie moléculaire (la réparation de l'ADN) et son succès dans certains pays comme les États-Unis ou les Pays-Bas (Partie 3) ;
- les manifestations spatiales de la résolution d'une question de recherche (la question du lien entre transcription et réparation de l'ADN) et les trajectoires individuelles de certains spécialistes de la réparation de l'ADN (Partie 3).

Sur la Figure 11, on montre comment ces objets se positionnent selon le niveau géographique et scientifique auxquels ils correspondent. Les numéros qui précèdent chaque objet indiquent l'ordre dans lequel ils sont abordés dans la thèse. Inspirée et nourrie de l'expérience du programme « Géoscience », l'approche dynamique, multi-niveau et « "trans-"quantitative et qualitative » – l'intérêt de ce dernier type d'approche est discuté dans les introductions au numéro double du *Bulletin de Méthodologie Sociologique* (Bidart & Dupray, 2014, 2015) – privilégiée s'est révélée précieuse pour cerner la science comme objet conceptuel ou plutôt les sciences comme réalités sociales, et montrer que ces réalités, d'abord appréhendées par le biais des publications scientifiques, sont à la fois le jouet des carrières individuelles et celui de tendances lourdes présentant d'importantes inerties.



Une approche multi-niveaux, quanti-quali et dynamique

Position des objets analysés dans la thèse numérotés par ordre d'apparition

Conception et réalisation: Marion Maisonobe

Figure 11 – Une approche multi-niveaux, quanti-quali et dynamique. Position des objets analysés dans la thèse numérotés par ordre d'apparition

Les effets d'inertie sont fondamentaux pour comprendre la spatialisation des activités scientifiques. C'est notamment l'avis de Michel Grossetti, qui s'en est expliqué le mardi 20 mai 2014 à l'occasion d'une séance du séminaire IBIFA à Toulouse. Ce séminaire IBIFA (pour Inside BIG Facilities) est coordonné depuis 2014 par le sociologue des organisations Vincent Simoulin, auteur d'un ouvrage sur le synchrotron de Grenoble (Simoulin, 2012) et Jérôme Lamy, spécialiste de l'histoire de l'observatoire de Toulouse (Lamy, 2007). IBIFA entend contribuer à structurer un objet d'études peu exploré par la littérature scientifique : les grands équipements scientifiques²² (Lamy & Simoulin, 2015). En effet, s'il existe plusieurs études de cas sur les grands équipements, il n'existe pas encore de trame qui en faciliterait l'étude et permettrait la comparaison. Voici donc le discours tenu par Grossetti le jour de la séance portant sur « les grands équipements et leurs territoires » : « Les spécialisations scientifiques peuvent avoir des effets sur le long terme y compris sur l'économie (...) Il y a une géographie de l'activité scientifique qui ne se réduit pas à une géographie économique ou à une géographie des ressources naturelles, et cette géographie ne peut se comprendre qu'en étudiant des processus historiques, l'installation ou la présence dans des lieux de données ou d'instruments (...) Il y a des effets d'inertie ou de "dépendance du sentier" (...), et on ne peut pas faire autrement que de faire l'histoire des activités pour en comprendre les logiques de répartition. Le monde scientifique est lent avec beaucoup d'inertie car les carrières sont longues et la mobilité institutionnelle est faible. Contrairement aux idées toutes faites qui voient les chercheurs comme des cerveaux nomades, la majorité des carrières se font entre un, deux ou trois lieux : il y a des liens forts dans ce milieu entre l'espace et le temps. ». Revenant sur sa propre trajectoire, il précise : « Pour ma part, je suis parti de la question : "Pourquoi y-a-t-il autant d'informaticiens à Toulouse ?". Cela supposait d'aller enquêter du côté de l'histoire des formations et du tissu économique toulousain. J'ai trouvé que nous pouvions découpler la question de la localisation des activités scientifiques de celle des activités économiques (même s'il y a toujours des recoupements) ; et qu'il existe un encastrement territorial, non seulement des activités mais aussi des savoirs (un chercheur peut mourir mais ces idées peuvent rester vivantes localement et continuer d'influencer les pratiques scientifiques) ». Cette dernière remarque suggère que ce n'est pas parce qu'on adopte une entrée institutionnelle qu'il faut rejeter la tradition anglo-saxonne mettant l'accent sur les circulations et les savoirs. Au contraire, il convient de travailler sur l'articulation des cinq éléments qui, d'après Grossetti, sont au cœur du rapport de la science au temps et à

²² D'après Simoulin et Lamy : « Aussi bien la mise en ordre de l'histoire longue des grands équipements scientifiques que leurs analyses sociologiques restent en effet à opérer » (voir la description figurant sur le site internet du Labex SMS : <http://sms.univ-tlse2.fr/accueil-sms/le-labex-sms/operations-structurantes/mondes-scientifiques>)

l'espace : les individus et leurs trajectoires, les styles cognitifs, les objets d'étude, les instruments et pour finir les institutions. On retrouve, formulés sous une forme plus personnelle, les éléments du tableau ci-dessous, obtenu à l'issue d'une réflexion collective menée au cours des réunions plénières du programme « Géoscience » en 2011 (Tableau 3).

Rapport à l'espace	Localisation	Liens	Mobilités
Entités			
« <i>Objets scientifiques</i> »	Lien avec les lieux (encastrement spatial)	Réseaux de textes, d'objets, etc.	Circulation
<i>Individus</i>	Distribution spatiale	Réseaux interpersonnels	Mobilités individuelles temporaires ou plus durables
<i>Organisations</i>	Distribution spatiale	Réseaux d'organisations	Création ou déplacement

Tableau 3 – *Les entités scientifiques et leur rapport à l'espace. Tableau repris d'un document de travail réalisé par Michel Grossetti et André Grelon (Compte rendu des réunions « Géoscience » des 11 et 12 mai 2011)*

Ce point de vue est important pour notre travail car il résume brièvement la teneur du projet de géographie des sciences sur lequel se fonde le précédent état de la littérature. En particulier, il conforte ce que nous souhaitons souligner au préalable, c'est-à-dire qu'il est nécessaire de jongler avec différents niveaux d'analyse et différentes temporalités : celles des carrières, des instruments, des idées et objets d'étude, des disciplines et des organisations, y compris économiques et politiques. Surtout, parce qu'il est formulé par l'un des coordinateurs de « Géoscience », ce point de vue permet de mieux saisir le déroulement du programme et son organisation (Figure 10). En effet, une grande part du programme « Géoscience » a été consacrée à l'histoire des systèmes nationaux d'enseignement supérieur et de recherche. Car, pour comprendre la répartition actuelle des activités scientifiques, il est apparu judicieux de tirer des fils pouvant nous faire remonter à des événements vieux de plusieurs siècles. En particulier, l'abolition des universités médiévales à la Révolution française s'est traduite par une refonte du système académique national sous Napoléon dont la géographie actuelle de la recherche et de l'enseignement supérieur est largement l'héritière. Saisir les manifestations géographiques de cette histoire nationale était l'objectif de la « tâche 1 » du programme « Géoscience », coordonnée par André Grelon, spécialiste des écoles d'ingénieur françaises du XIX^e siècle.

La première mission des membres de « Géoscience » a été de reconstituer les étapes de la construction du système scientifique français depuis la Révolution française.

Cette mission a rassemblé des historiens français, pour la plupart issus de l'histoire des sciences et des techniques et de l'histoire de l'enseignement supérieur. Nombre d'entre eux étaient spécialistes d'une zone géographique particulière, zone dans laquelle ils ont été amenés à éplucher les archives d'institutions locales, de facultés et/ou d'écoles (par exemple Caroline Barrera pour la faculté de Toulouse, Yves Lequin pour l'Institut universitaire de technologie de Belfort-Montbéliard, Françoise Birck pour l'École des mines de Nancy etc.). Puisqu'il fédérait des spécialistes des systèmes locaux d'enseignement et de recherche de plusieurs villes françaises, ce chantier s'inscrivait dans la continuité du programme PIR-Villes sur les villes et les institutions scientifiques, déjà évoqué dans le Chapitre 3.2.1. En revanche, l'ouvrage obtenu à l'issue de ce chantier présentera la particularité d'adopter une approche résolument spatiale (Grossetti & Grelon, à paraître). Il proposera des cartes inédites montrant l'évolution de la répartition des activités scientifiques en France au cours des deux derniers siècles, et analysera l'articulation des logiques nationales d'organisation de la carte universitaire aux logiques locales — qui relèvent des luttes d'influence des élus, des élites intellectuelles et entrepreneuriales ou encore des effets de site et de position géographique. L'introduction de l'ouvrage, réalisée par René Sigrist, situera la France dans un contexte plus large : la fondation de systèmes scientifiques nationaux à l'échelle européenne ; tandis que pour aborder la période contemporaine, le dernier chapitre s'appuiera sur les données du WoS, traitées dans le cadre de la « tâche 2 » du programme « Géoscience ».

Preuve que l'histoire a pris une place importante pendant le déroulement du programme, il a été un peu plus tard décidé qu'un second ouvrage serait coordonné au sein de la première tâche, sur l'histoire de plusieurs systèmes nationaux et sur les dynamiques géographiques des activités scientifiques à l'échelle mondiale (Kleiche-Dray, à paraître). Cette deuxième mission a mis à contribution des historiens, des géographes et des spécialistes des « Pays en Développement » (PED) ou « non-hégémoniques » (cette appellation se veut plus neutre, comme l'indiquent Losego & Arvanitis, 2008). Parmi ces spécialistes, un certain nombre venait de l'ancienne équipe PED de l'ORSTOM, désormais intégrée à l'axe « Éducation et savoirs au Sud » du CePeD²³ (Centre Population et Développement). Leur participation à cet ouvrage socio-historique a permis de retracer l'émergence de plusieurs « communautés nationales de recherche ». L'expression « communauté nationale » étant satisfaisante dans ce cadre car elle revêt une dimension quasi-programmatique. Issue de l'idéologie du cercle mertonien (voir Chapitre 4.1), elle a été relayée par les Organisations Non Gouvernementales ou internationales comme l'UNESCO, puis récupérée par les élites intellectuelles locales pour promouvoir le déve-

²³ Le CePeD est une unité mixte regroupant l'Université Paris Descartes, l'Institut de Recherche sur le Développement (IRD) et l'INED (l'Institut National d'Études Statistiques).

loppement de la recherche et l'enseignement supérieur dans des territoires anciennement colonisés et pour garantir leur autonomie vis-à-vis du « centre ». Pour ce second ouvrage, l'information spatiale contenue dans le WoS est également exploitée, mais cette fois-ci pour fournir des données de cadrage sur la production scientifique mondiale et son évolution par grandes aires géographiques. Aussi, réalisée par nos soins, la représentation des données de collaborations scientifiques entre villes russes (Figure 12) devrait accompagner le chapitre portant sur les évolutions du système scientifique depuis le règne de Pierre Le Grand à nos jours (Eckert, Gouzévitch, Gouzévitch, & Pane, 2014).

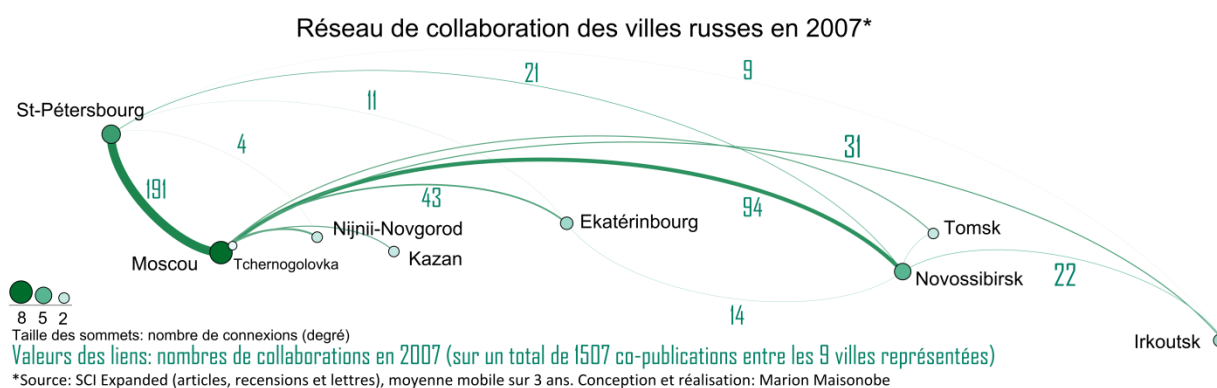


Figure 12 – Réseau de collaborations scientifiques des agglomérations russes en 2007.

Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

En plus d'apporter des compléments et des éléments de démonstration tout au long de l'ouvrage, les données scientométriques viendront alimenter un chapitre de synthèse, auquel nous avons contribué, qui résume les grandes tendances mondiales capturées par le biais du WoS (Eckert, Grossetti, Jégou, & Maisonobe, 2014). Il s'agit des grands résultats obtenus dans le cadre de la « tâche 2 » sur la géographie de la production et des collaborations scientifique contemporaine, lesquels sont également restitués dans la deuxième partie de la thèse (Chapitre 9).

Contrairement aux autres tâches (1 et 3) qui mettaient en relation des chercheurs géographiquement dispersés, la « tâche 2 » a été principalement menée à Toulouse à deux exceptions près : Myriam Baron, géographe, qui a contribué au chantier depuis Paris, et Yves Gingras, historien des sciences, qui a apporté le savoir-faire et l'aide de son équipe spécialisée dans les traitements bibliométriques depuis Montréal. Bien plus que les autres tâches, la « tâche 2 » présentait la particularité de demander des efforts et des capacités importantes en matière de traitement quantitatif. En contrepoint, dans le cadre de la « tâche 3 » chargée des études de cas, l'étude de la mobilité des étudiants entre Grenoble

et la Colombie reposait sur un travail à mi-chemin entre le qualitatif et le quantitatif dont se sont occupés les sociologues des sciences Dominique Vinck et Gloria Zarama ²⁴.

Avant d'en venir au principe et au déroulement de la « tâche 2 », retenons que la dimension quantitative n'était qu'une facette du programme de recherche et au-delà du projet de géographie des sciences qui s'y trouve attaché. Relevant de ce même esprit, la présente thèse n'est pas limitée à une dimension quantitative. En particulier, nous verrons à travers une étude de cas portant sur une spécialité scientifique en biologie moléculaire comment les résultats obtenus à partir des traitements quantitatifs peuvent être utilement complétés et enrichis par des approfondissements d'ordre qualitatif. Cette démonstration arrivera au moment où l'on étudiera la géographie de la communauté de la Réparation de l'ADN (Partie 3).

1.2. Une analyse quantitative de l'évolution de la répartition spatiale des publications scientifiques et des collaborations (tâche 2)

Pour les coordinateurs du programme de recherche, la « tâche 2 » devait permettre d'approfondir et de vérifier des hypothèses formulées à l'issue d'études de cas réalisées au préalable, notamment celles de Michel Grossetti et Béatrice Milard portant sur la « territorialisation de la recherche » dans plusieurs pays (Grossetti & Losego, 2003 ; Milard & Grossetti, 2006 ; Grossetti *et al.*, 2009). En particulier, il s'agissait d'étendre à tous les pays et aux années les plus récentes la recherche des effets de « régionalisation » et d'« internationalisation », et par conséquent de déterminer, à travers la répartition spatiale de la production scientifique, si le phénomène de déconcentration des activités scientifiques déjà observé au sein de plusieurs pays (France, Espagne, Portugal, Afrique du Sud, Russie) pouvait être considéré comme un phénomène global.

Mais pour faire une géographie de la production scientifique de façon systématique, il n'était pas possible de se limiter à des extractions ponctuelles de notices bibliographiques à partir des bases de données en ligne. Cette fois-ci, il fallait un accès plus complet à la base privilégiée jusque-là pour des études de cas sur des villes ou des pays en

²⁴ En 2012, leur travail a été présenté au XIX^e congrès international des sociologues de langue française à Rabat :

Vinck, D. (2012, juillet). Emmener ou pas les problématiques de son pays lorsqu'on est un doctorant en mobilité. Présenté à Formation à la recherche et Enseignement Technique et Scientifique, XIX^e congrès international des sociologues de langue française à Rabat.

Zarama, G. (2012, juillet). Les doctorants étrangers entre deux eaux. Création de liens entre groupes de recherche du Nord et du Sud via la mobilité des doctorants. Présenté à Formation à la recherche et Enseignement Technique et Scientifique, XIX^e congrès international des sociologues de langue française à Rabat.

particulier : le *Web of Science* (WoS). Cette source s'est imposée parce qu'elle est la référence mondiale parmi les spécialistes de l'information bibliographique, et comme il est possible d'en consulter des extraits depuis les ressources en ligne du CNRS (*via* BiblioSHS, ce qui n'est pas le cas de *Scopus*), elle était déjà familière à plusieurs d'entre nous. En fait, le catalogue de références connu sous le nom de WoS, depuis qu'il a été mis en ligne sur le *World Wide Web*, comprend le plus ancien index de citations automatisé : le *Science Citation Index* (SCI). Datant des années 1960, c'est aussi – comme on le verra dans la seconde partie de ce chapitre – l'index le plus complet ou représentatif de la production scientifique mondiale. Il a d'ailleurs servi de modèle aux index concurrents constitués plus récemment, comme *Scopus* lancé en 2004 par l'éditeur scientifique privé Elsevier.

Dès lors, le traitement de l'information géographique contenue dans le WoS était le projet central de la « tâche 2 » du programme « Géoscience ». Plus précisément, il fallait mettre au point une méthode permettant d'étudier la géographie des activités scientifiques à partir du contenu des adresses institutionnelles renseignées pour chaque notice bibliographique, c'est-à-dire pour chaque entrée dans la base de publications. Surtout, la conviction selon laquelle les données de production scientifique devaient être analysées à un niveau géographique plus fin que celui des pays afin de saisir des tendances locales et leurs articulations aux tendances globales (ce qui n'avait jamais été fait auparavant !) faisait l'originalité de ce projet et l'inscrivait définitivement dans la mouvance du courant de « scientométrie spatiale » impulsé depuis les Pays-Bas à la fin des années 2010, et évoqué dans le Chapitre 4.2.

Pour que ce projet se réalise, il fallait se tourner vers des organismes disposant de la base de données tels que les organismes gouvernementaux de statistiques publiques spécialisés dans le traitement de l'information sur les sciences et les techniques. De tels organismes existent à Paris et à Montréal depuis les années 1990.

À Paris, l'Observatoire des Sciences et Techniques a été créé au début des années 1990 — près de 10 ans après la création du réseau universitaire des Unités Régionales de Formation à l'Information Scientifique et Technique (URFIST). La nécessité d'un organisme national spécialisé dans le traitement de l'information scientifique s'est affirmée tout au long des années 1980. On la trouve présente dès 1983 dans le rapport de stage de Rigas Arvanitis, soutenu par Bruno Latour et Michel Callon, sur le système de construction et d'utilisation des indicateurs de production scientifique aux États-Unis (Arvanitis, 1983), qui est issu d'un intérêt pour le développement d'un système analogue en France, et elle est encore démontrée dans l'ouvrage de synthèse de Madeleine Brocard sur les enjeux français en matière d'aménagement de la recherche (Brocard, 1991). À la fin de son ouvrage, Brocard retrace ainsi le processus ayant conduit à la création de l'OST depuis 1985, au moment où a été faite la proposition d'un « observatoire national du développe-

ment de la recherche et de la technologie dans les régions » qui aurait pour mission de « tenir à jour l'état et l'évolution du potentiel de recherche publique et industrielle par région sur le territoire français » jusqu'au rapport « demandé par un nouveau ministre en 1988, qui aboutissait... enfin à la décision de mettre en place un Observatoire des Sciences et des techniques (OST) sous la forme d'un Groupement d'Intérêt Public constitué par les Ministères concernés, l'INSEE et quelques partenaires privés. » (*ibid.*, p. 221). Résumant les missions initialement dévolues à l'organisme, Brocard explique : « L'OST devrait être animé par une cellule centrale légère d'une dizaine de personnes associées à de nombreux partenaires français et étrangers, et chargée de produire des indicateurs, d'assurer leur diffusion, d'améliorer progressivement leur qualité. » (*loc. cit.*). En 1991, bien qu'elle soit soulagée de savoir que cet organisme a enfin vu le jour, Brocard regrette deux choses : d'une part qu'il soit essentiellement tourné vers l'élaboration d'indicateurs destinés à élaborer des politiques nationales au détriment d'indicateurs qui pourraient servir à prendre des décisions plus locales et pas nécessairement politiques ; et d'autre part que les régions ne soient pas associées à cet observatoire (*ibid.*, p. 221-222). Pour elle, ces limites sont la preuve que :

« La prise de conscience n'est pas encore faite, parmi les administrateurs de l'État, que ce qui intéresse la nation peut aussi intéresser les régions, et que la seule façon d'assurer la cohérence de la politique nationale en matière de recherche, tant réclamée par les gestionnaires de la science, c'est d'associer tous ceux qui doivent y participer financièrement. » (*ibid.*, p. 222). Désormais, si l'OST s'intéresse à la fabrication d'indicateurs pouvant servir aux régions, voire aux métropoles, sa mission reste nationale et s'accompagne d'une dimension européenne, puisque depuis que Ghislaine Filliatreau en a pris la direction en 2005, cette mission est définie comme suit : « l'OST est chargé de fournir les indicateurs de la mission interministérielle pour la recherche et l'enseignement supérieur (MIREs), qui permettent un suivi des différents programmes gouvernementaux dans ce domaine (production scientifique, participation à la construction de l'espace européen de la recherche) »²⁵. Enfin, en 2014, il a été décidé que l'OST s'associe au nouvel institut national d'évaluation de la recherche, le Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (HCERES) (Pumain & Dardel, 2014). En fait, dans la même logique qu'au Royaume-Uni ou en Italie (Rebora & Turri, 2013), il est dorénavant question, et c'est là un point-clef des récentes réformes des pays européens, de compléter l'évaluation par les pairs ou « à dire d'expert », qui est coûteuse à faire appliquer régulièrement et susceptible de générer des conflits d'intérêt, par les informations suppo-

²⁵ Source: URL : http://fr.wikipedia.org/wiki/Observatoire_des_sciences_et_des_techniques ; <http://www.obs-ost.fr/>

sées plus objectives qu'apportent les indicateurs scientométriques (Ploux-Chillès & Heintz, 2012 ; Sousa & Brennan, 2014).

Bien que son acronyme soit le même, l'Observatoire des Sciences et Technologies (OST) de Montréal a un statut plus académique, et par conséquent des intérêts sensiblement différents. Abrité par l'Université du Québec à Montréal (UQAM), il s'agit d'une « unité de soutien à la recherche du Centre interuniversitaire de recherche sur la science et la technologie (CIRST) »²⁶. Il est actuellement dirigé par Yves Gingras qui est non seulement historien des sciences, mais aussi un éminent spécialiste de la sociologie des sciences et de la scientométrie. En 1997, Gingras a cofondé l'OST à Montréal avec un autre historien, Benoît Godin, déjà évoqué en Chapitre 2.1, qui s'en est détaché au cours des années 2000.

Pour des raisons de proximité organisationnelle et cognitive, notre cellule de travail toulousaine comprenant Denis Eckert, Michel Grossetti, Laurent Jégou, Béatrice Milard et moi-même a bénéficié de la collaboration de l'OST de Montréal dès le commencement de la tâche 2. Cela tenait au fait que le CIRST participait officiellement au programme « Géoscience », mais aussi qu'Yves Gingras connaissait déjà Michel Grossetti et Béatrice Milard depuis plusieurs années. Cette collaboration nous a permis d'entrer assez tôt en contact avec l'univers de la scientométrie. Entre Toulouse et Montréal, un partage des tâches s'est progressivement instauré.

Lors d'une réunion à Montréal en Octobre 2011, les équipes ont convenu que le travail de localisation et de codage géographique de la production scientifique à partir des adresses de publications serait réalisé à Toulouse et que Montréal apporterait son savoir-faire en matière de scientométrie, et plus particulièrement de mesure des publications (bibliométrie). Ainsi, pendant plus d'un an, notre équipe s'est concentrée sur l'opération de géocodage de la table d'appartenance géographique associée aux informations stockées par Montréal sur la production scientifique.

Plus particulièrement et dans la perspective d'étudier des évolutions, l'équipe montréalaise nous a transmis le contenu de trois champs permettant de localiser les publications du *Science Citation Index Expanded* (SCI Expanded) à 4 années différentes (1978, 1988, 1998 et 2008) : un champ « ville » (la localité géographique telle qu'elle est indiquée par les auteurs ayant publié pendant ces années-là), un champ « province » (les états ou provinces dans le cas des pays fédéraux) et un champ « pays ». Chaque combinaison unique de ces trois champs, est appelée « triplet » sachant que plusieurs triplets peuvent servir à désigner la même localité puisqu'il existe différentes façons de déclarer une même appartenance géographique.

²⁶ Source: URL : <http://www.ost.uqam.ca/fr-ca/observatoire/présentationdelost.aspx>

À l'issue du travail de codage géographique – on reviendra sur la procédure ultérieurement, dans la Partie 2.Chapitre 6.1.3 – notre équipe est parvenue à associer une unique localité à chaque triplet, ainsi qu'un unique jeu de coordonnées géographiques (latitude, longitude) pour chaque localité participant à des publications scientifiques. Le résultat de ce premier travail collectif a été consigné dans un « fichier recette » c'est-à-dire un fichier qui, pour chaque « triplet-source » indique un « triplet-destination » proposant une seule graphie et un seul jeu de coordonnées par localité publiante. Comme le suggère la Figure ci-dessous (Figure 13), ce fichier est utile pour structurer les données bibliographiques en vue d'une analyse spatiale multi-niveaux.

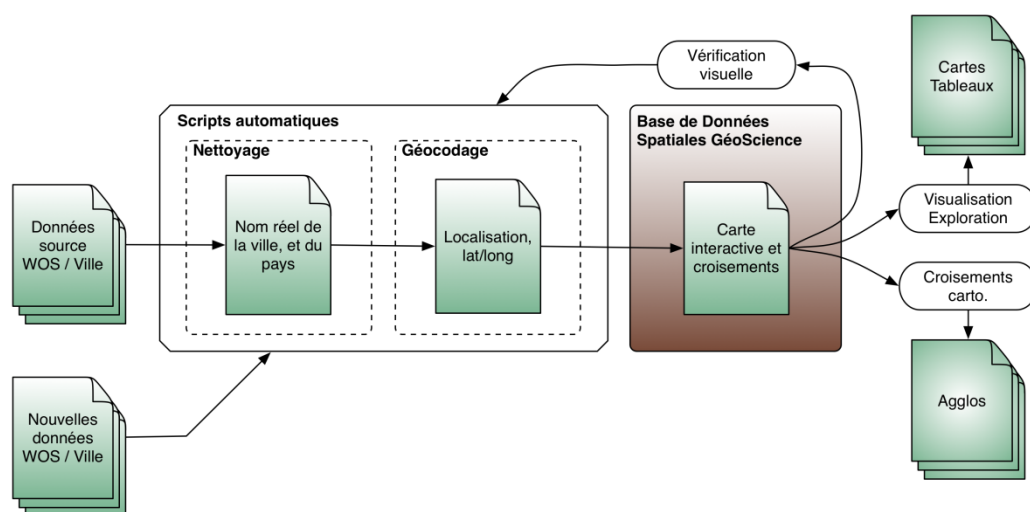


Figure 13 – Du géocodage à l'analyse spatiale des données bibliographiques. Schéma extrait d'une présentation de Laurent Jégou, Montréal, octobre 2011

Parallèlement à ce travail de codage, un second chantier a été conduit par Denis Eckert, Myriam Baron et Laurent Jégou pour identifier des ensembles ou agglomérations de localités. Pour procéder à des analyses spatiales à l'échelle mondiale, il a été décidé de regrouper l'information au niveau d'entités spatiales locales délimitées à l'aide de critères unifiés, tels que la densité de population ou l'artificialisation du sol. Sans cet effort d'harmonisation des entités élémentaires d'analyse, il aurait fallu se fier aux découpages administratifs municipaux alors qu'ils reposent sur des critères qui varient d'un pays à l'autre. À l'aide d'une méthode en deux étapes, décrite dans le numéro spécial de la revue *M@ppemonde* et détaillée plus loin, les 18654 localités repérées à partir de la simplification et du codage des adresses ont pu être regroupées au sein de 10729 agglomérations (Eckert, Baron, & Jégou, 2013). C'est grâce à cette procédure que des localités comme Villejuif ou Palaiseau ont été intégrées à l'entité parisienne dont la production scientifique aurait autrement été éclatée en une multitude de petites communes.

Suite au codage géographique des localités et à la délimitation d'agglomérations susceptibles d'accueillir les publications, il ne restait plus qu'à mesurer et analyser l'information. Les données de publications du SCI Expanded pour les 4 dates retenues (1978, 1988, 1998 et 2008) ont été « ventilées » par agglomérations et par pays avec l'aide d'Yves Gingras et Vincent Larivière de l'OST Montréal. Ce travail a permis de vérifier et prolonger les résultats préalables de Michel Grossetti et Béatrice Milard et ainsi de confirmer l'hypothèse d'une déconcentration de la répartition des publications à différents niveaux spatiaux. En effet, les analyses ont montré qu'à chaque date, les publications scientifiques ont été produites par un nombre de plus en plus important de pays, et à l'intérieur de ces pays, par un nombre de plus important d'agglomérations urbaines. Ces résultats ont donné lieu à un article collectif paru dans la revue *Urban Studies* en 2014 (Grossetti *et al.*, 2014). L'article va à contre-sens des travaux, inspirés de l'économie géographique, présumant que la mondialisation serait à l'origine d'un inévitable mouvement de concentration de l'activité scientifique dans les espaces métropolitains (voir Chapitre 3.2.2).

Au terme de cette première recherche, nous disposons d'un « fichier recette » permettant de « nettoyer » et coder l'information géographique du SCI Expanded à un niveau de précision inégalé auparavant. Un tel fichier présentait une certaine valeur pour un organisme comme l'OST Paris habitué à fournir de l'information soit au niveau des pays, soit au niveau des provinces ou régions en s'appuyant sur les découpages administratifs. En particulier, pour les analyses à l'échelle européenne, l'organisme n'avait d'autres choix que de recourir à la nomenclature NUTS (*Nomenclature of Territorial Units for Statistics*), qui propose des équivalences entre niveaux de découpages administratifs d'un pays à l'autre, on y reviendra. En comparaison, notre approche présentait l'originalité de ne dépendre d'aucun découpage administratif et ainsi, de permettre la comparaison à l'échelle mondiale entre entités urbaines définies selon un critère unifié. L'intérêt de l'OST Paris pour cette façon d'aborder l'information spatiale a donné lieu à une rencontre dans leurs locaux de la rue Vaugirard à la fin de l'année 2012. Cette réunion a permis que s'établisse un nouveau partenariat.

À l'OST Paris, la possibilité de disposer directement de l'information bibliographique du WoS à plusieurs dates s'est présentée. Une fois familiarisée avec les enjeux de la bibliométrie, notre équipe a jugé intéressant d'héberger une partie de la base à Toulouse. La flexibilité serait alors plus grande pour réaliser des extractions, et des requêtes, mais aussi pour se charger du traitement des données de collaborations en construisant des matrices de co-signatures entre lieux. La seule condition à respecter était de ne pas en faire un usage commercial. Cette étape de la tâche a été l'occasion d'ouvrir la boîte noire de la bibliométrie. Progressivement, l'observatoire parisien nous a délivré l'ensemble des enre-

gistements du WoS à sa disposition, à savoir les tables de publications, de revues et d'appartenance géographique pour les 3 index du WoS (le SCI Expanded, l'AHCI et le SSCI) et pour toutes les années à partir de 1999 (1999-2012). Pour stocker cette base de données sur le serveur de l'Université de Toulouse, le système *open-source* de management relationnel PostgreSQL a été sélectionné²⁷. D'après Laurent Jégou, « PostgreSQL avait deux caractéristiques pertinentes pour notre travail : une bonne adaptation aux applications en ligne, et une très bonne gestion des données géographiques à travers son extension PostGIS. PostGIS permet de gérer la composante spatiale des données et d'exécuter des requêtes spatiales en langage SQL. Il permet un accès et un traitement rapide d'une grande quantité de données. De plus, l'utilisation d'un système relationnel (sur un serveur *web*) a permis sans trop de mal de créer des applications de géovisualisation, qui se sont révélées très adaptées pour notre grande équipe de chercheurs dispersés. Cela a offert des moyens d'explorer et d'évaluer la progression du géocodage. Enfin, en utilisant un système de scripts, on a pu documenter et automatiser toutes les opérations de transformation des données, depuis le stade de l'adresse brute à celui du point de publications précisément géolocalisé. »²⁸ (Jégou, 2014).

C'est en grande partie sur les choix méthodologiques et les traitements effectués à partir de ces nouveaux jeux de données que se fonde la seconde partie de la thèse. Dans le cadre de l'opération Netscience du Labex SMS qui a débuté en 2013, le travail d'équipe s'est poursuivi autour de ces nouvelles données. Avec Laurent Jégou, j'ai participé à définir et construire des matrices de co-publications entre lieux pour les années 1999-2001 et 2006-2008, et à observer comment se manifestaient les phénomènes simultanés de « régionalisation » et d'« internationalisation » préalablement identifiés par Michel Grossetti et Béatrice Milard. Il est apparu à cette occasion que le phénomène d'internationalisation n'était pas partagé de manière équivalente par tous les pays et que le niveau national restait un niveau extrêmement structurant pour l'organisation des collaborations scientifiques. Actuellement, les résultats de ces traitements sont en cours de publications et de nou-

²⁷ Source : URL : <http://www.postgresql.org>

²⁸ L'original : « PostgreSQL has two relevant characteristics for our work: a perfect integration in online applications contexts, and a very good handling of geographical data through its spatial extension PostGIS. PostGIS offers the capacity to manage the spatial components of data as database fields and to execute spatial queries in SQL language. It provides not only a fast access to a very big quantity of data but also a very adapted one, taking into account our need of spatial processing at the core of the data management. Moreover, the use of a network database system (on a web server) enabled the creation of a handful of geovisualization applications, which were a helpful asset to our geographically dispersed team of researchers. It offered means to explore and evaluate the progression of the geocoding, while under way. Finally, by using a scriptable database system, we have been able to document and script each successive operation of data transformation, from original addresses character strings to geographical points precisely defined. » (Jégou, 2014).

veaux traitements sont en voie de réalisation permettant d'étudier les variations par disciplines et d'étendre les analyses jusqu'à l'année 2012. Parallèlement, et à titre plus personnel, j'ai obtenu auprès de l'OST Paris l'ensemble des enregistrements du WoS utilisant le mot-clef « *DNA repair* ». Dans le cadre d'une étude de cas sur la spécialité de la réparation de l'ADN, cela a permis d'explorer les données du WoS rassemblant l'ensemble des travaux des biologistes de cette spécialité sur la période contemporaine. Une partie de ces données a été exploitée pour l'article publié dans la revue *DNA Repair* en collaboration avec une biologiste spécialisée dans ce domaine (Maisonobe, Giglia-Mari, & Eckert, 2013).

En somme, la seconde étape du travail d'équipe nous a amenés à travailler autrement pour deux raisons :

- d'une part parce que nous avons d'abord travaillé avec l'ensemble des données du WoS alors que les travaux avec l'OST Montréal portaient essentiellement sur le SCI Expanded, qui est une extension du SCI, l'index de citations le plus ancien et le plus reconnu dont la vaste ouverture (élargie dans les années 1990) va de la biomédecine aux mathématiques en passant par la physique, la chimie, les sciences de l'univers ou encore, les sciences de l'ingénieur (le SCI Expanded). Or, comme on va le voir, le WoS rassemble deux autres index de citations importants : le SSCI (*Social Science Citation Index*) pour les sciences sociales et l'AHCI (*Arts & Humanities Citation Index*) pour les humanités. Ces deux bases, le SSCI et l'AHCI, sont moins anciennes et moins fiables d'après les avis d'experts mais nous les avons stockées et traitées pour ensuite avoir la liberté de les intégrer ou les exclure des analyses ;

- et d'autre part parce que, cette fois-ci, il n'était pas possible de remonter au-delà de 1999. Cette contrainte amputait largement la profondeur historique de nos résultats par rapport aux traitements qui avaient pu être réalisés en collaboration avec l'OST Montréal. Cela dit cette seconde limite a sans doute permis d'échapper à la tentation de comparer sur une base semblable des situations très hétérogènes et d'en tirer de mauvaises interprétations. En effet, plus on remonte dans le temps, plus on prend de risques méthodologiques car la comparabilité des indicateurs bibliométriques est moins pertinente pour au moins quatre raisons.

Premièrement, les pratiques de publication ont changé, en particulier depuis les années 1990, puisque la publication d'articles et notamment les co-signatures d'articles ont été largement encouragées (Pontille, 2000). À côté des incitations à l'obtention d'un meilleur impact et d'une plus grande efficacité scientifique, il y a des effets de croissance propres à l'augmentation de la force de travail scientifique et à l'augmentation de la littérature scientifique associée. Dès lors, il est délicat de faire la distinction entre les effets de croissance réelle et les effets qui tiennent à l'évolution des pratiques de recherche.

Ce problème explique l'importance des enjeux de normalisation statistique pour les scientomètres. Par exemple, Olle Persson *et al.* ont essayé de mesurer le rôle de l'augmentation générale des co-signatures sur le reste des indicateurs scientométriques, et notamment le nombre de références par articles tous domaines confondus. De leur analyse, nous avons extrait le graphique suivant qui estime l'intensité des évolutions d'ensemble intervenues simultanément au cours de la dernière génération (Figure 14).

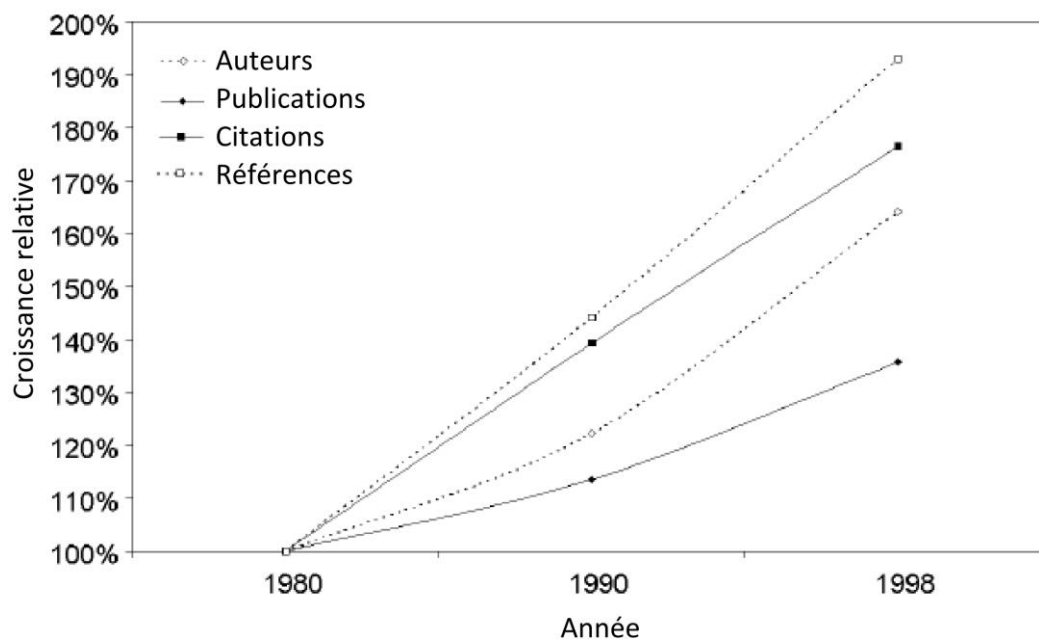


Figure 14 – Croissance relative des publications, des auteurs, des références et des citations (tous domaines confondus). Extraite et traduite de l'article « *Inflationary bibliometric values : The role of scientific collaboration and the need for relative indicators in evaluative studies* » (Persson, Glänzel, & Danell, 2004).

Deuxièmement, l'organisation spatiale n'est pas la même, et outre l'apparition de nouveaux pays résultant, pour une grande partie, de l'éclatement du bloc soviétique, l'étendue des agglomérations urbaines s'est modifiée ce qui rend les comparaisons statistiques entre entités géographiques moins robuste à travers le temps, on y reviendra (Julien, 2000).

Troisièmement, la qualité d'indexation des publications dans la base n'est pas stable dans le temps. Comme nous le verrons à l'occasion de l'étude de cas sur la spécialité de la réparation de l'ADN, il n'est pas rare, parmi les notices de publications antérieures à 1990, que les adresses ne soient pas renseignées ou que seulement une seule adresse soit enregistrée, y compris quand il s'agit très formellement d'une co-publication.

Et enfin, quatrièmement, la couverture des index, c'est-à-dire le périmètre des revues couvertes, s'est considérablement élargi et modifié année après année. Et puisque nous avons jugé ce phénomène essentiel pour cette thèse et avons constaté qu'il était peu étudié, comparé par exemple à l'évolution des pratiques de publications, nous allons examiner en détail les déterminants de la sélection des revues et de l'évolution du WoS dans la section suivante.

Il faut dire que l'activité de publication d'articles s'est développée et le nombre de scientifiques et de lieux de sciences n'a cessé d'augmenter au cours du XX^e siècle, ce qu'a tendance à traduire la croissance du périmètre des index bibliographiques. À notre sens, cette croissance est, en elle-même, un objet très pertinent à étudier pour saisir les tendances de fond au niveau mondial. Par exemple, en regardant de plus près la couverture des revues russes et celle des revues chinoises, il est possible d'en déduire les dynamiques à l'œuvre – y compris en se passant des nombres d'articles proprement dits qui sont, en fait, corrélés au nombre de numéros parus et de revues indexées par unité d'analyse (coefficient de corrélation = 0,87, p -value = 2,2e-16 pour l'année 2008 au niveau des pays). Cette approche suppose que l'on considère la base elle-même et l'évolution de son contenu comme un laboratoire de l'état de l'activité scientifique dans le monde.

Pour déterminer si cette conception du WoS est acceptable, la section suivante revient sur les origines et la nature de cette source. Cette étape est importante avant de discuter des différents moyens de traiter l'information géographique associée aux données bibliographiques, de l'analyser et d'en tirer des conclusions concernant l'évolution de la géographie des activités scientifiques dans le monde. En effet, il faut être conscient que parce que nous avons affaire à une source privée dont le siège est aux États-Unis, le « monde scientifique » auquel elle donne accès n'est pas neutre. Ce « monde » et sa logique sont même le résultat d'une vision enchantée, la vision du créateur de l'ISI, du SCI, du SSCI et de l'AHCI, Eugène Garfield.

2. Une source tertiaire: le *Web of Science*

« La Toile, Lieu de la communication : La Toile qui semble avoir bouleversé le monde en connectant tous les lieux "virtuellement", présente plusieurs aspects contradictoires de la mondialisation. Loin d'en permettre la transparence à tous, elle reproduit l'inégalité héritée qui vaut pour toutes les innovations technologiques au moment de leur insertion sociale. Mais plus sérieusement sans doute, elle se propose comme un lieu d'attente à tous ceux qui souhaitent s'exposer au monde. » (Retaillé, 2012, p. 73)

2.1. « L'ère romantique » de l'Institute for Scientific Information (ISI) ou les origines du Web of Science (WoS)

La source tertiaire qui fait autorité parmi les spécialistes des études portant sur la mesure de la science est : le *Web of Science* (WoS). Il comprend l'un des plus anciens index de publications académiques automatisés destinés à la récupération d'information (ou *information retrieval*) : le *Science Citation Index* (SCI). Chaque notice de publication indexée renvoie aux notices des documents référencés par la publication ainsi qu'aux notices des documents qui l'ont citée *a posteriori*. L'idée d'un tel système apparaît pour la première fois dans un écrit de Vannevar Bush, un peu moins connu que *Science, the Endless Frontier*, qui est *As We May Think* (Bush, 1945). Il y décrit « Memex », un dispositif technique idéal pour la recherche d'information.

Bon nombre de causes plaident en faveur du développement d'index de publications scientifiques aux États-Unis dans les années 1950 parmi lesquelles : la nécessité apparue en temps de guerre d'être en mesure de s'informer rapidement sur les avancées scientifiques aux contenus opérationnels ; les nouvelles possibilités offertes par l'informatique, alors à l'état embryonnaire, pour faciliter le stockage, la récupération et la circulation d'informations ; le besoin des chercheurs de situer leur travaux dans un univers de travaux existants au-delà des frontières disciplinaires et nationales, et de citer leurs sources ; enfin, le sentiment d'un retard par rapport aux Soviétiques en matière de management scientifique et de systèmes d'information (Garfield & Hayne, 1955 ; Garfield, 2006). Somme toute, dans un contexte de foisonnement de la production scientifique, la mise au point d'un tel outil se révèle indispensable aux chercheurs et aux administrateurs américains de la science pour s'y retrouver. Ce besoin est comblé dans les années 1960 par le linguiste

Eugene Garfield dans le cadre d'un institut privé, l'*Institute for Scientific Information* (ISI) à Philadelphie. Garfield s'inspire du principe du *Sheppard's Citation*, un dispositif datant de 1873 qui recense les délibérations judiciaires pour créer le *Science Citation Index* (SCI). Il est disponible à l'achat sous forme imprimée dès 1963, quelques années plus tard sur bandes magnétiques, et plus tard encore sur CD-ROM.

Avec l'intention de couvrir l'ensemble des disciplines scientifiques, Eugene Garfield et son équipe complètent le SCI par deux autres index dans les années 1970 : le *Social Science Citation Index* (SSCI) et l'*Art and Humanities Citation Index* (AHCI). Depuis les années 1990 et le rachat de l'ISI par l'entreprise multinationale Thomson Reuters (qui marque la fin de ce que certains scientomètres ont qualifié « d'ère romantique » de l'institut), de nouveaux index ont été ajoutés parmi lesquels le *Conference Proceedings Citation Index* et le *Book Citation Index*. Consultable en ligne sur internet (l'accès n'est pas gratuit), la combinaison de ces index est dorénavant appelée le WoS. La particularité des index de citations est qu'en naviguant de références en références par le biais des citations, il n'y a plus à se préoccuper des frontières disciplinaires. Le périmètre de la recherche n'est pas limité, comme avec un index traditionnel, à des thématiques de recherche cloisonnées.

Les services de l'ISI bénéficient peu à peu d'une grande notoriété : ils sont mobilisés par les chercheurs en général pour leurs recherches bibliographiques et pour mesurer la visibilité de leurs publications ; par les instances d'administration de l'activité scientifique de différents pays, organismes et laboratoires pour le calcul d'indicateurs servant à l'évaluation de la recherche : le service des relations internationales de l'ISI joue un rôle crucial en ce sens dès les années 1960 (Garfield, 1972) ; par les chercheurs en *science studies* pour alimenter leurs études ; ou encore par les documentalistes ou spécialistes de l'information. Il ne serait pas exagéré d'affirmer qu'en de nombreux points, leur popularité et leur utilisation dépasse ce que les fondateurs de la société avaient pu anticiper. Assez tôt, la visibilité d'une revue ou d'un auteur (mesurée par le nombre de citations reçues), au moins en biomédecine et aux États-Unis, apparaît influencée par le fait d'être ou de ne pas être recensé par l'ISI. Comme il décide quel est le contenu scientifique qu'il indexe et dont il permet la circulation, l'institut américain en vient à disposer d'une influence considérable sur le « champ scientifique ». Évidemment, son niveau d'influence dépend du crédit que les chercheurs lui accordent et de son périmètre de diffusion, tous deux amenés à varier selon les disciplines et les espaces géographiques. Mais, il reste dans l'intérêt de l'ISI de prétendre disposer d'une large audience pour convaincre un maximum d'éditeurs scientifiques et d'organismes scientifiques de s'abonner²⁹. Ainsi, pour des disciplines

²⁹ En fait, d'après David Pontille et Didier Torny, il semblerait que Garfield ait d'abord rencontré des difficultés à amortir le coût du SCI. La survie de l'entreprise de Garfield aurait alors été assurée par les profits générés par les *Current Contents* (Pontille & Torny, 2013).

comme la médecine, la biologie, la chimie ou encore la physique, être recensé par l'ISI est indiscutablement valorisant, mais on ne peut toujours pas en dire autant des mathématiques ou des sciences sociales, en France notamment.

L'institut joue un rôle autrement important dans l'émergence de la scientométrie. À l'occasion de la création de la revue *Scientometrics* en 1978, Eugene Garfield se souvient : « On a toujours su que le *Citation Index* pourrait servir pour des travaux d'historiographie et de sociologie. Cependant, on n'avait jamais imaginé que l'accès à des index de citations à grande échelle pourrait stimuler le développement d'un domaine entièrement nouveau, la scientométrie. »³⁰ (Garfield, 1978). Il faut dire qu'au cours de sa carrière, Eugene Garfield travaillait de près avec Derek de Solla Price qui est considéré comme le père fondateur de la spécialité³¹. Ce dernier définissait à l'origine la scientométrie comme la science de tout ce qui, concernant l'activité scientifique, peut être mesuré, mais les travaux publiés dans la revue *Scientometrics* sont en majorité des travaux de bibliométrie (autour de la littérature scientifique : parutions, citations, auteurs, adresses institutionnelles) exploitant des données de l'ISI (Polanco, 1995). Au bout du compte, étant donné la variété des services proposés par l'ISI depuis sa création, qui vont du *Web of Knowledge* (qui englobe davantage de sources que le *Web of Science*), au *Web of Science* qui regroupe les principaux index, en passant par le *Journal Citation Report* (JCR), ainsi que, des années durant, les *Current Contents* (CC) qui sont des veilles bibliographiques hebdomadaires (dont la couverture ne recouvre pas exactement celle des index de citations) à partir desquelles l'ISI établissait l'*Author Address Directory* et les *Weekly Index*, il existe énormément de façons différentes de mobiliser les données de l'institut. En outre, il est délicat de comparer les différents résultats obtenus à partir de ces données parce que les méthodes de comptages peuvent varier, qu'il n'existe pas de consensus parmi les scientomètres sur la méthode à privilégier (Gaufriau, Larsen, Maye, Roulin-Perriard, & Ins, 2008), on y reviendra, mais aussi et surtout parce que la couverture ou liste des revues couvertes par les différents services n'a pas cessé d'évoluer.

La couverture des bases n'a jamais eu vocation à être exhaustive et Eugene Garfield est souvent revenu dans ses billets intitulés *Current Comments* (publiés conjointement aux

³⁰ L'original : « When the Citation Index was first developed, its primary purpose was information retrieval. We were always aware of the possible uses for historiographical and sociological research [Garfield; Historiographs, Librarianship, and the History of Science, in *Toward a Theory of Librarianship: Papers in Honor of Jesse Hauk Shera*, C. H. Rawski (Ed.), Scarecrow Press, Metuchen, N. J. 1973.]. However, we never dreamed that the availability of large scale citation indexes would spur the development of an entirely new field of scientometrics. » (Garfield, 1978).

³¹ Eugene Garfield écrit en 1987 : « So when he [Robert K. Merton alias Bob] called me one day to suggest that we honor our mutual friend, Derek Price, by a joint effort, I was thrilled - indeed, exhilarated - by the opportunity to join my mentor in this special tribute ». *Essays of an information scientists*, vol. 10, p. 72, 1987 (Garfield, 1987).

Current Contents) sur le mode de sélection des revues en revendiquant comme objectif la couverture de revues scientifiques de qualité. Il prend conscience progressivement de disposer d'un pouvoir qu'il n'avait pas mesuré à l'origine si bien qu'il s'écrie en 1973 : « Je n'aime pas le pouvoir qui m'a été donné, ou imputé, de "contrôler" la destinée de revues, en particulier les petites et nouvelles revues, en raison du succès de CC. Cela me donne des responsabilités qui ne sont pas les bienvenues. »³² (Garfield, 1973b). En effet, le périmètre des index et de *Current Contents* s'est révélé décisif à trois égards au moins : premièrement en raison des enjeux de visibilité voire de reconnaissance scientifique pour les éditeurs et les auteurs en lien avec l'évaluation scientifique, deuxièmement en raison de l'influence du périmètre sur la science elle-même puisque les lectures des chercheurs ainsi que leur appréciation de la littérature en dépendent, et troisièmement en raison de l'influence du périmètre sur les mesures qui servent à l'étude sociale des activités scientifiques en *science studies*. Pour ces trois raisons, le mode de sélection des revues provoque des critiques et des interrogations multiples, les plus récurrentes étant le biais anglo-saxon de l'ISI et le biais en faveur de la biomédecine.

Comme la biomédecine a servi de modèle à l'élaboration du premier index de citations, c'est à elle que les méthodes bibliométriques et mesures popularisées par l'ISI sont le mieux adaptées. C'est d'ailleurs ce dont témoigne le mode de calcul du facteur d'impact des revues³³. Dans un numéro de *Current Comments* paru en novembre 1979, Garfield raconte : « Notre politique de sélection s'est développée au fil du temps à mesure que CC et nos autres services se sont épanouis. Quand la première version de CC/*Life*

³² L'original : « I do not like the power that has been vested, or imputed, to me to "control" the destiny of journals, especially small and new one, because of CC's success. It imposes an unwelcome responsibility. » (Garfield, 1973b).

³³ Le facteur d'impact est un indicateur qui dispose d'une grande popularité dans l'univers de la recherche scientifique. Pouvant être calculé de nombreuses façons différentes, il doit rendre compte du différentiel de visibilité des travaux de recherche relevant de leur support de publication. Certaines revues sont plus lues que d'autres et par conséquent, les articles ou notes qui y paraissent ont davantage de chance d'être cités. L'interprétation du facteur d'impact est variable d'une discipline à une autre : s'il est rare qu'un article en biomédecine ne soit pas cité dans les deux ans qui suivent sa parution, c'est beaucoup plus fréquent pour un article de mathématiques. Les bibliomètres ont consacré beaucoup d'énergie à affiner le calcul du facteur d'impact : il peut se calculer à 2 ou à 5 ans et il existe des grilles permettant de l'interpréter correctement selon l'ordre de grandeur spécifique à la discipline voire à la spécialité concernée. En France, la cellule bibliométrique de l'INSERM utilise des grilles très précises pour évaluer le travail de ses équipes qui tient compte des variabilités intra-disciplinaires. Cf. : Haeffner-Cavaillon, N. (2012, février). La bibliométrie comme outil d'évaluation et de pilotage à l'INSERM. Présenté à Groupe de réflexion sur l'analyse et la caractérisation de la production scientifique-Outils et Méthodes, Urfist de Paris.

Notes consultables à l'adresse <http://urfist.enc.sorbonne.fr/veille-et-recherche/ateliers/evaluation-caracterisation-bibliometrie/seance-du-12-fevrier-2012>

Ajoutons que dans ces travaux, Yves Gingras montre les limites de cet indicateur et rappelle qu'il est conçu pour mesurer la visibilité d'une revue, et non celle d'une équipe, encore moins celle d'un chercheur (Gingras, 2008).

Science [rendant compte sur une base hebdomadaire des parutions en sciences de la vie] est parue en 1957, nous n'avions aucune règle formelle pour décider quels journaux devaient être inclus. (...) Puisque nous ne couvrons que 200 revues à cette période, nous nous limitons aux meilleures. »³⁴ (Garfield, 1979). La couverture s'élargit progressivement à partir de la création du *Science Citation Index* qui permet de repérer les revues les plus citées par les revues déjà recensées. Pour l'intégration de nouvelles revues, l'ISI est attentif au comité éditorial de la revue, à l'impact scientifique de ses membres, à celui des revues dans lesquels ils publient (mesuré par le biais des citations reçues), ainsi qu'à des avis d'experts (*ibid.*). Évidemment, le problème avec ce système c'est celui du serpent qui se mord la queue : être recensé par l'ISI augmente la probabilité d'être cité et les revues recensées ont donc plus de chances d'être citées, à qualité équivalente que les revues qui ne bénéficient pas de la même visibilité.

D'après Garfield, le problème de la sélection des revues se pose véritablement pour les jeunes revues ou les revues de second rang et non pour le « haut du panier ». Or, puisque les services de l'ISI sont payants, il y a une clientèle à satisfaire ce qui influence le choix des revues : « Quand nous entamons l'évaluation de revues "moins importantes", on est obligé de considérer d'autres facteurs. L'un des plus importants est l'intérêt de nos lecteurs : par exemple, les biochimistes constituent une part considérable de notre lectorat. Ainsi, une nouvelle revue en biochimie ou en biologie moléculaire sera prioritaire sur une revue d'horticulture, ou sur une revue publiée par une société médicale locale. »³⁵ (Garfield, 1973b, *op. cit.*). Suivant cette logique commerciale, l'ISI tend à entretenir ses biais initiaux en faveur de certaines spécialités scientifiques comme la biologie moléculaire ou plus généralement les sciences naturelles.

Garfield insiste aussi sur la forme : ce n'est pas qu'une question de contenus, il faut que les parutions soient régulières, que l'adresse des auteurs soit bien renseignée, de même que le titre et le résumé pour faciliter l'opération de scannage. La question de la forme est une manière commode de justifier le biais anglo-saxon des index car c'est sur le modèle des publications américaines que l'opération de scannage est conçue (Garfield, 1979, *op. cit.*).

³⁴ L'original : « Our selection policies developed over time as CC and our other information services grew. When the forerunner of *CC/Life Sciences* first appeared in 1957, we had no formal policies to decide which journals should be included. (...) Since we covered only 200 journals at that time, we limited ourselves to the best. » (Garfield, 1979).

³⁵ L'original : « When we start evaluation of "less important" journals, we are forced to consider other factors. One of the most important is reader interest, e.g., biochemists compose a major segment of our readership. Therefore, a new journal in biochemistry or molecular biology will get a higher priority than a journal in horticulture, or one published by a local medical society. » (Garfield, 1973b, *op. cit.*).

De plus, intégrer des titres est coûteux et Garfield reconnaît que l'attitude plus ou moins positive et volontariste de certains éditeurs peut aussi intervenir dans la sélection. Dans les années 1970, sous la direction de Garfield, l'ISI ne refuse pas de subventions provenant des éditeurs souhaitant que telle ou telle parution de second rang soit recensée du moment qu'elle répond aux critères de sélection les plus élémentaires. Si après un certain temps, la revue intégrée n'est pas suffisamment citée, l'ISI peut décider de ne plus la recenser. Cela-dit, la volonté de Garfield n'est pas de contribuer ou d'encourager l'inflation de la production scientifique, bien au contraire : à ses yeux, limiter le nombre de revues recensées est un moyen d'entretenir la compétition scientifique et au-delà la qualité de la production (Garfield, 1973b, *op. cit.*).

Dès l'origine, certains éditeurs se sont montrés très méfiants vis-à-vis du travail de l'ISI. En particulier, certains ont vu d'un mauvais œil, comme une concurrence à leurs propres services commerciaux, le service baptisé OATS (*Original Article Tear Sheet*) mettant à disposition de clients (bibliothèques ou individus) qui en feraient la demande (*reprint requests*), la version réimprimée d'un article recensé dans CC auquel ils ne peuvent pas accéder (Garfield, 1979, *op. cit.*). Garfield s'en explique à plusieurs reprises dans *Current Comments* en soulignant combien ce service est utile, notamment pour les pays en voie de développement qui n'ont pas, à proximité, de bibliothèques et services de documentation richement fournis (Garfield, 1970b). Une autre façon de se procurer des tirés-à-part consiste bien sûr à contacter directement les auteurs, ce que l'ISI favorise aussi à travers *Current Contents*. Systématiquement, les adresses des premiers auteurs sont indiquées dans les veilles hebdomadaires pour faciliter le contact entre les scientifiques et pour leur permettre d'échanger directement les versions réimprimées de leurs travaux par voie postale (Garfield, 1970a). L'esprit des services proposés par l'institut démontre bien que Garfield envisageait l'ISI comme un facilitateur de la circulation des idées à l'échelle planétaire.

À ce titre, il se faisait régulièrement le prophète de la langue anglaise en incitant les éditeurs des pays en développement, les éditeurs français y compris, à faire publier, au moins leurs résumés, en anglais (Garfield, 1967). Il en est même venu à exprimer le vœu que disparaisse l'alphabet cyrillique et à s'étonner que Lénine ne se soit pas chargé de le faire disparaître quand il en avait le pouvoir (Garfield, 1975). Pour lui, la standardisation de l'information est nécessaire à sa circulation et rend possible l'internationalisme nécessaire au progrès scientifique (Garfield, 1973a). Il s'amuse de se faire régulièrement taxer d'« impérialiste » par ceux qui, comme les scientifiques français, sont réticents à utiliser la langue anglaise (Garfield, 1977). Il parvient même à s'attirer les foudres de la France et des pays francophones à l'occasion de ce qu'Yves Gingras qualifie d'« affaire Garfield », c'est-à-dire les protestations et débats, y compris à l'Académie des sciences de Paris, qui

ont suivi la publication en 1976 par Garfield dans la *Recherche* d'un article provocateur intitulé « La science française est-elle trop provinciale ? » (Gingras, 2002).

L'idéologie de l'internationalisme scientifique est manifeste dans le discours de Garfield et, au-delà, dans la stratégie d'expansion de l'institut. L'ISI entend élargir géographiquement sa clientèle pour ne pas rester un service américain. En 1972, Garfield déclare : « pour que des services d'information scientifique soient optimaux, ils doivent être aussi internationaux que la science elle-même. »³⁶ (Garfield, 1972). D'ailleurs, en plus des actions visant à promouvoir ses services à l'étranger, notamment auprès des agences gouvernementales, l'ISI tient compte de la géographie pour la sélection des revues couvertes. Garfield l'évoque en 1979, dans l'un de ses billets déjà cité sur la sélection des revues : « Nous cherchons à savoir à quel point le sujet du journal est important pour les lecteurs de CC, et quel pourcentage du lectorat va être intéressé par cette revue. Nous estimons l'actuelle couverture que nous avons de cette thématique. Est-ce que notre couverture actuelle est suffisante ? Au-delà, nous évaluons notre couverture de l'aire géopolitique représentée par ce journal. »³⁷ (Garfield, 1979, *op. cit.*). De la même manière que l'ISI cherche à s'ajuster aux intérêts scientifiques de ses clients, il peut décider du sens dans lequel élargir la couverture de la base en se fondant sur leur origine géographique. De ce fait, au mépris d'une bonne représentativité de la distribution des activités scientifique, l'ISI a sans doute choisi d'entretenir certains biais géographiques pour satisfaire sa clientèle — de la même manière que, du point de vue des thématiques scientifiques, il a d'abord favorisé la biochimie et la biologie moléculaire.

En 1973, un peu moins d'un client de l'ISI sur deux réside en dehors des États-Unis (Garfield, 1973b, *op. cit.*). Un indicateur intéressant du lectorat de l'ISI à ses débuts est donné par l'origine des demandes de réimpressions (*reprint requests*). Celles-ci sont examinées dans un article plusieurs fois cité par Garfield : « *The Geography of Reprint Requests* » de Solomon et Jennifer Sarah Posen, professeurs de médecine à l'Hôpital de Sydney (Posen & Posen, 1969). Les Posen analysent en détail les 2500 demandes de réimpression qu'ils ont reçues directement entre mars 1967 et juillet 1968 pour sept de leurs articles. Pour trois de ces articles soit 482 demandes, ils regardent de près si la forme de l'adresse correspond à la forme publiée dans *Current Contents* ou bien à la forme qui figure sur l'article original. Pour 40 % des demandes, l'adresse à laquelle est adressée la de-

³⁶ L'original : « Our corporate planning has recognized and is building on the fact that optimum science information services must be as international in input, application, and feedback as science itself. » (Garfield, 1972).

³⁷ L'original : « We consider how important the subject area of the journal is to readers of CC, and what percentage of that readership is going to be interested in the journal. We assess the depth of our present coverage of the field. Is our present coverage sufficient? Further, we weigh our coverage of the geopolitical area represented by the journal. » (Garfield, 1979, *op. cit.*).

mande est reprise de *Current Contents*, ce qui est loin d'être négligeable. Alors que le quart des demandes totales (128 demandes sur 482) ne provient pas d'Amérique du Nord, les demandes qui ne proviennent pas d'Amérique du Nord représentent près de 30 % des demandes qui sont calquées sur le modèle de *Current Contents* (56 demandes sur 192). Ce résultat, bien que difficilement généralisable, semble témoigner du fait que *Current Contents* bénéficie d'une audience qui ne se cantonne pas à l'Amérique du Nord. Mais, il serait délicat d'en dire davantage sur la distribution géographique de *Current Contents* à partir de ce seul cas de figure (Tableau 4) :

L'adresse des auteurs est mentionnée...	Amérique du Nord	Europe non communiste	Europe communiste	Amérique Latine	Asie	Afrique	Australie	Total
...comme sur l'article	177	38	7	9	7	1	1	240
...comme dans <i>Current Contents</i>	136	33	10	5	8			192
...autrement	40	6	1	1	1			50
Total	353	77	18	15	16	1	1	482

Tableau 4 – Origine des demandes de réimpressions des articles de Posen et Posen. Extrait et traduit de l'article « *The Geography of Reprint Requests* » (Posen & Posen, 1969)

Pour résumer, il semble qu'Eugene Garfield a laissé une forte empreinte sur les services qu'il a contribué à créer et que les billets d'humeur publiés pendant sa carrière rendent bien compte de l'esprit qu'il cherchait à donner à son « œuvre ». Son idéologie est centrale parce que les choix qui ont été faits des années 1960 aux années 1990 concernant la sélection des revues ont une influence pratiquement irréversible sur les mesures de la science disponibles pour cette époque.

En effet, s'il arrive à l'ISI, depuis son rachat par Thomson Reuters, d'enrichir le contenu des années passées, cette démarche reste limitée à quelques « initiatives ponctuelles ». C'est en tout cas ce que nous a laissé entendre Guillaume Rivalle, responsable de clientèle chez Thomson Reuters à l'occasion d'un échange de courriels datant de mars 2012 : « En général, la couverture s'élargit mais en se tournant vers le futur. Lorsqu'une revue est ajoutée au WoS, les publications de celle-ci sont indexées à partir de cette date d'ajout uniquement car c'est à cette date que l'on a estimé la revue comme étant "de qualité suffisante" pour apparaître dans la base. Il arrive cependant que nous élargissions la couverture d'un index entier, par exemple l'index des sciences sociales allait auparavant jusqu'à 1956 et nous l'avons étendu en une fois à 1900. Ce sont des initiatives ponctuelles. Récemment nous avons créé le *Book Citation Index* (index de citations des ouvrages) et nous

y indexons des livres dont la parution remonte jusqu'à 2005. »³⁸. De ce fait, les choix réalisés par le passé restent déterminants pour comprendre le contenu de la base. Pour nous aider à mieux connaître ce contenu, Guillaume Rivalle nous a gracieusement fait parvenir un fichier détaillant l'année d'inclusion et le cas échéant d'exclusion de 17 302 revues couvertes par les services de l'ISI. Pour chacune des revues, le fichier nous indique leur durée d'indexation dans l'un (ou plusieurs) des index du WoS entre 1899 et 2008, certaines revues pouvant être recensées dans plusieurs index à la fois. Nous en avons extrait les courbes suivantes fondées sur le contenu actuel des enregistrements du WoS (le nombre de revues indexées par an). Premièrement, la Figure 15 décrit l'état en 2008 de la couverture du WoS (soit avant l'extension du SSCI évoquée dans la citation) :

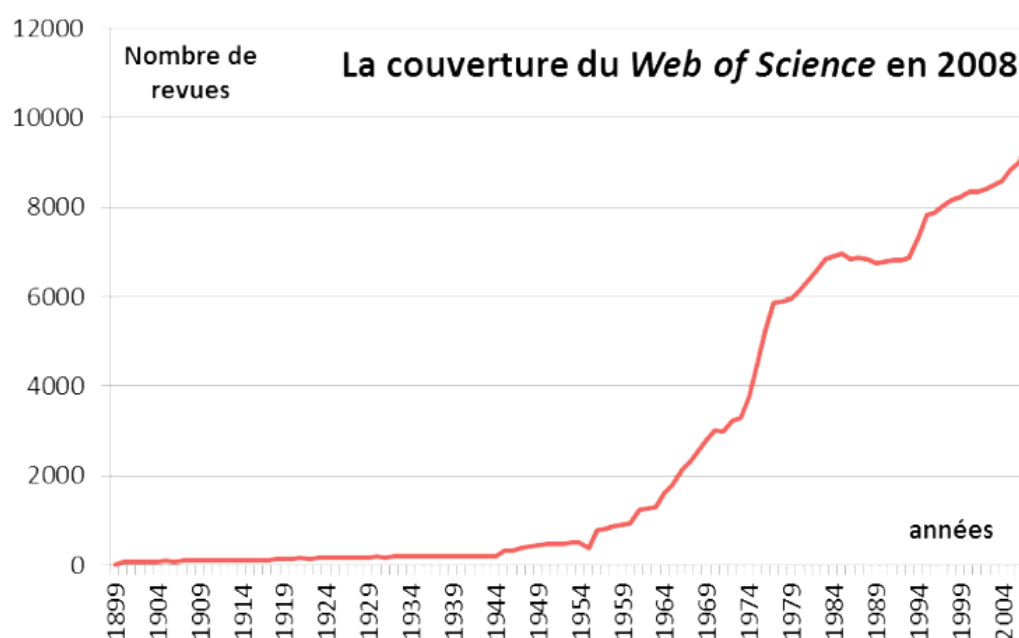


Figure 15 – La couverture du Web of Science en 2008 : comptage du nombre de revues indexées par an. Source : ISI Thomson Reuters

Deuxièmement, la Figure 16 permet de voir le détail de la couverture du SCI Expanded, du SSCI et de l'AHCI (sachant que quelques revues sont comprises dans plusieurs index à la fois).

³⁸ Courriel reçu le 17 avril 2012 en provenance de Guillaume Rivalle. Objet : phd géographie des activités scientifiques.

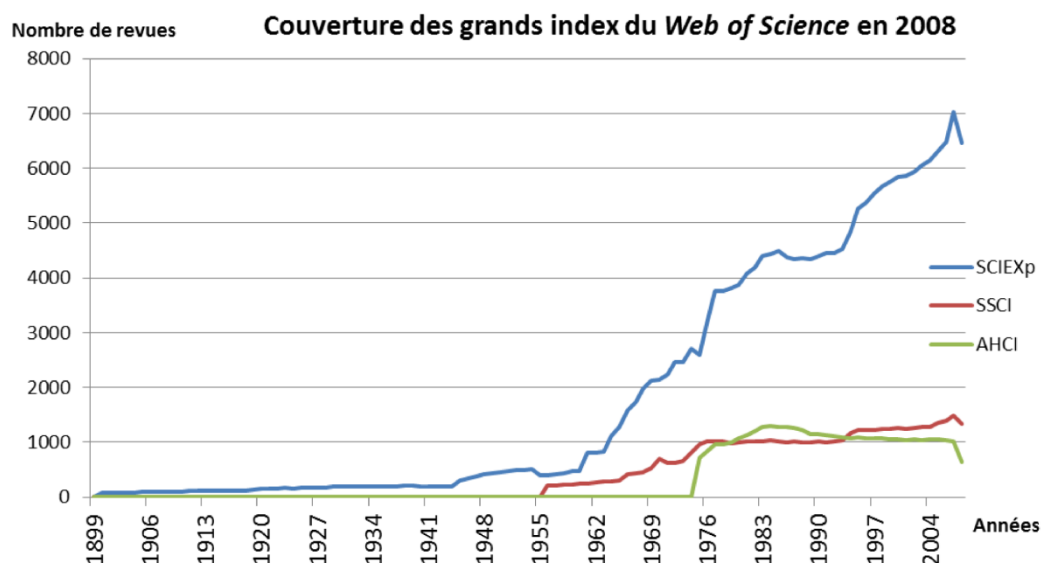


Figure 16 – La couverture des grands index du Web of Science en 2008 : comptage du nombre de revues indexées par an et par index. Source : ISI Thomson Reuters

Ces deux graphiques montrent combien les périmètres des index de citations varient dans le temps. Sachant que seule une petite partie des revues existantes dans le monde est couverte, il est difficile de savoir dans quelle mesure l'évolution observée sur les graphiques reflète l'évolution générale du nombre réel des revues scientifiques ou si elle dépend des choix, de la politique ou des possibilités de l'institut de Philadelphie. Les éditoriaux d'Eugene Garfield témoignent bien du fait qu'il était conscient des limites et des biais de la base dont il était l'artisan. Il entendait toutefois défendre une ligne et une idée. Pour lui, cet outil allait pouvoir faciliter la circulation de la science en rendant visible les travaux scientifiques par-delà les frontières disciplinaires et géographiques et, de cela, il était à la fois très fier et enthousiaste. La base qu'il a contribué à créer est longtemps restée inégalée en nombre de revues couvertes tous pays et disciplines confondus, ce qui de l'avis de nombreux spécialistes en scientométrie en fait un outil de choix pour rendre compte de la production scientifique au niveau mondial et de son évolution. Soulignons que les critères de sélection actuels des revues sont décrits dans un document de James Testa traduit en plusieurs langues et accessible en ligne sur le site de Thomson Reuters (Testa, 2012).

2.2. Les données de l'ISI : reflet de la science mondiale ?

À la fin des années 1970, lorsque Tibor Braun lance la revue *Scientometrics*, les services bibliométriques de l'*Institute for Scientific Information* (ISI) ne connaissent aucun équi-

valent en matière de visibilité et de nombre de revues scientifiques couvertes. Partant de ce constat, les spécialistes des sciences de l'information ou scientomètres s'appliquent à rendre compte de l'état de la production scientifique dans le monde à partir des données de l'ISI. Ces données sont même jugées plus adaptées que celles des organismes de statistique publique pour les comparaisons internationales. Par exemple, au lieu d'utiliser les données trop hétérogènes d'organismes nationaux ou internationaux (OCDE, UNESCO), András Schubert et Tibor Braun préfèrent estimer la force de travail scientifique à partir d'un comptage s'appuyant sur le nombre d'auteurs dont les publications sont indexées dans le SCI : « L'utilisation d'une base de données indépendante et homogène pour estimer la force de travail scientifique (mondiale) semble être à privilégier »³⁹ (Schubert & Braun, 1992). Un extrait du tableau comparatif qu'ils obtiennent est reproduit ci-dessous (Tableau 5). Ces deux chercheurs sont ceux qui ont fondé la première unité de recherche consacrée à l'analyse de citations l'*Information Science and Scientometrics Research Unit* (ISSRU) à Budapest en 1978 et qui ont lancé, au cours des années 1980, avec Wolfgang Glänzel, une rubrique intitulée *Word Flash of Basic Research* dans la revue *Scientometrics*, donnant régulièrement le détail, par pays, des indicateurs de production scientifique calculés à partir du traitement dans leur laboratoire des bandes magnétiques de l'ISI.

Pays	Nombre de premiers auteurs	Potentiel de publication	Estimation de l'UNESCO
Algérie	213	1012,7	n.a.
Argentine	2934	10926,6	9500
Bangladesh	209	1345,6	n.a.
Brésil	4252	31692	24015
Chili	1473	6150,8	n.a.
Colombie	186	1025,5	1449
Costa-Rica	127	818,1	320
Egypte	2151	6575,3	n.a.
Ethiopie	108	685,1	n.a.
Ghana	125	434,2	4084
Hong Kong	676	2476,2	n.a.
Inde	17585	41890,6	28233
Indonésie	157	—	764
Iran	309	—	n.a.
Irak	454	6444	n.a.

Tableau 5 — Le SCI versus les estimations de l'UNESCO. Tableau extrait et traduit de l'article « *Three scientometric studies on developing countries as a tribute to Michael Moravcsik* » (Schubert & Braun, 1992)

³⁹ L'original : « The use of an independent, common database for estimating scientific manpower seems rather advisable » (Schubert & Braun, 1992).

Si les spécialistes des sciences de l'information ont très tôt ventilé les données scientométriques par pays, cette pratique n'est devenue systématique qu'avec la création de cette rubrique dans la revue *Scientometrics*. En effet, dans leurs états de la littérature sur ce thème K.C. Garg et Benoît Godin évoquent plusieurs chercheurs ayant pris l'initiative de mesurer la distribution géographique des hommes de sciences entre la fin du XIX^e siècle et le milieu du XX^e siècle (Garg, 2003 ; Godin, 2006a). Mais c'est l'ISI qui a rendu possible la systématisation de ces traitements, sans jamais échapper cependant au développement de la concurrence. L'état de l'art de Garg, qui va jusqu'à la fin des années 1990, mentionne donc des travaux empiriques fondés sur d'autres bases que celles de l'ISI. Il existe par exemple les bases PASCAL et FRANCIS que les spécialistes français de l'étude des sciences dans les pays « non-hégémoniques » ont eu tendance à privilégier pendant un temps (Arvanitis, Waast, & Gaillard, 2000). Le scepticisme de ces spécialistes au sujet des données de l'ISI n'est pas étonnant compte tenu de la faible couverture des revues non anglophones. Cette dernière, nous l'avons évoqué, puise son origine dans la certitude d'Eugene Garfield concernant la supériorité scientifique des États-Unis, à quoi il faut ajouter la localisation de l'institut à Philadelphie. Malgré cette conviction, lorsque, dans les années 1980, le débat s'ouvre au sujet de la représentativité des pays en développement ou non-anglophones, la revue *Scientometrics* sert d'espace pour la polémique et Garfield accepte volontiers d'y participer.

À cette occasion, le sort réservé aux pays sous-représentés est le cheval de bataille de Michael J. Moravcsik, physicien se préoccupant dès les années 1960 du développement de la science dans les pays du « Tiers-Monde » (Moravcsik, 1966). Il défend dans les années 1980 la création de bases bibliographiques vouées à couvrir la science produite dans les aires géographiques les plus mal représentées par l'ISI. Il l'évoque comme un enjeu important pour le développement de la science dans ces territoires : « Pour des raisons qui sont très légitimes, il y a peu de chance qu'un organisme comme l'ISI puisse gérer la couverture des milliers de revues publiées dans le Tiers-Monde. J'ai pour cela suggéré depuis longtemps à diverses personnes concernées dans les pays en développement que ces pays, seuls ou en groupes, initient leur propre compilation informatisée de revues, auteurs, publications et peut-être même de citations. Un tel programme serait d'une grande valeur non seulement dans un contexte international, mais aussi en matière de politiques nationales et régionales, et pour étudier en science de la science les efforts fournis par les pays en développement, autrement dit pour évaluer et estimer les efforts dans ces pays, etc. »⁴⁰

⁴⁰ L'original : « I have therefore been suggesting for some time to various appropriate people in the developing countries that these countries, singly or in groups, initiate their own computerized compilation of journals, authors, publications and perhaps even citations. Such a program would be of great value not only in an international context, but also in terms of national and regional science policies, in terms of studies in

(Blickenstaff & Moravcsik, 1982). Cette déclaration suggère que ce n'est pas dans les moyens de l'ISI, ni dans son intérêt de couvrir des revues au « rayonnement local », mais que cet enjeu est celui des pays qui entendent conduire une véritable politique scientifique. Pour autant, Moravcsik ne renonce pas à l'idée d'améliorer la couverture de l'ISI. Il préside même le chantier de discussion ouvert par l'institut pour renforcer la couverture de la science produite dans les pays en développement. Lancé en 1985, cet effort collectif donne lieu à un rapport final dit « programme de Philadelphie » : *Strengthening the coverage of third world science* (Moravcsik, 1985). À cette époque, l'ISI estime couvrir entre 2 et 8 % des journaux du Tiers-Monde mais considère qu'il recense la moitié de la production de qualité issue des pays en développement (« *which meets the international standards of excellence* »). La discussion se concentre donc spécifiquement sur la légitimité et la possibilité de couvrir la moitié restante.

Le postulat très répandu parmi les scientomètres est qu'il est possible de distinguer entre une science « *mainstream* » ou « internationale » de qualité et une science « locale » ou « régionale » qui ne serait pas vouée à rayonner au-delà d'une aire géographique bien délimitée (Zitt, Perrot, & Barré, 1998). Le problème est qu'il y a parmi les revues indexées par l'ISI, bon nombre de revues que l'on pourrait qualifier de « locales ». C'est ainsi que pour une meilleure représentativité de la « science mondiale », il existe des voix pour proposer une solution inverse, à savoir qu'afin de couvrir les productions de la science d'excellence mondiale, il faudrait retirer des revues plutôt qu'en ajouter. C'est ce que le scientomètre français Michel Zitt appelle, avec Élise Bassecoulard, « le problème de Sivertsen » (du nom du premier scientomètre à avoir défendu cette position) par opposition au « problème de Moravcsik » (Zitt & Bassecoulard, 1999). Pour Zitt et Bassecoulard : « le choix de l'ISI (ou n'importe quelle autre base de données sélective) de sélectionner un journal ukrainien tourné sur l'Ukraine plutôt qu'un journal franco-français ou un journal coréen de portée coréenne, tous à impact facteur faible, est un choix très difficile, en particulier dès lors que son impact peut être affecté, entre autres facteurs, par la taille de la communauté scientifique correspondante et aussi par la question du langage plutôt que par la qualité intrinsèque du journal. » (*ibid.*). En accord avec la pensée de Garfield, ils affirment : « la sélection de la queue de distribution du *Science Citation Index* (en termes d'impact et/ou d'internationalisation) porte bien plus à controverse que la sélection des meilleures revues. » (*ibid.*). Ils en déduisent qu'il « est plus facile de résoudre ce qu'il appelle le problème de Sivertsen (réduire le périmètre du SCI) que le problème de

the science of science of efforts in the developing countries, in terms of evaluative, and assessing efforts in those countries, etc. » (Blickenstaff & Moravcsik, 1982).

Moravcsik (élargir le périmètre du SCI) qui est au cœur du "programme de Philadelphie". »⁴¹ (*ibid.*).

Zitt est sans doute l'un des scientomètres à avoir le mieux étudié les biais géographiques associés au périmètre du SCI. Alors qu'il n'existe pas d'indicateurs officiels pour mesurer le degré d'internationalité d'une revue, ni d'ailleurs pour lui affecter une nationalité (qui ne soit pas celle de son siège d'édition), il a proposé des méthodes de normalisation pour estimer le niveau d'internationalité compte tenu du domaine de recherche de la revue et à partir de la nationalité des auteurs contribuant à la revue. En 2003, il démontre avec ses collègues que l'impact et la part mondiale des pays varie énormément selon que l'on inclut ou pas (dans les calculs) les revues à forte orientation nationale et à faible impact (Zitt, Ramanana-Rahary, & Bassecoulard, 2003). Les résultats obtenus diffèrent notamment entre les grands pays et les petits pays. Les revues originaires de ces derniers sont plus rarement limitées à un lectorat local, contrairement aux revues originaires de grands pays comme la France ou la Russie qui, en plus, ne sont pas forcément en anglais. Donc, paradoxalement, élargir le périmètre tend à favoriser les petits pays au détriment des gros — notamment ceux qui comme les Pays-Bas ont une force éditoriale importante et sont tournés vers l'international. D'après Zitt, pour optimiser son image nationale, il vaut mieux avoir peu de revues couvertes mais avec une forte ouverture vers l'international et de ce fait un bon impact ; plutôt que beaucoup de revues couvertes avec une faible ouverture internationale et un faible impact. Les conclusions de Zitt sont plus mathématiques qu'idéologiques : il ne veut pas s'opposer à l'élargissement du périmètre des index, mais il veut démontrer que les indicateurs bibliométriques sont surtout pertinents pour comparer des revues qui présentent un profil semblable.

Or, l'élargissement du périmètre du SCI a tendance à ajouter une forte hétérogénéité à l'ensemble, ce qui complique l'interprétation des indicateurs. Il résume : « Aussi longtemps que les revues à comité de lecture sont un bon outil de communication scientifique, et par extension de description du paysage scientifique international dans la plupart des disciplines, il faut se souvenir que des représentations très différentes de la science

⁴¹ L'original : « the choice by ISI (or any selective database) of a national-oriented Ukrainian journal rather than a national-oriented Korean or French journal, all supposed with low-impact factors, is a very difficult one, particularly since impact may be affected, among other factors, by the size of the corresponding national scientific community and possibly by the language question, rather than the intrinsic quality of journals. Therefore, the lower tail of SCI (in terms of impact and/or internationalization) is far more controversial than its top journals collection (a census of the best sources). Perhaps, the "bigger" the science is, the safer the description using standard bibliometric tools. A further selection within the SCI may be envisaged for certain types of countries indicators, and is technically manageable. It is easier to solve what we could call the "Sivertsen's problem" (restricting SCI), than the "Moravcsik's problem" (expanding SCI) expressed in the "Philadelphia Program". But Sivertsen's proposal does not address the possible Anglo-American bias in "international" journals. » (Zitt & Bassecoulard, 1999).

mondiale peuvent être obtenues selon le choix de la base de données et le périmètre d'inclusion. »⁴² (*ibid.*). De leur côté, Thed N. van Leeuwen *et al.* avancent des arguments comparables à ceux de Zitt en étudiant l'évolution de l'anglais dans le périmètre du SCI (van Leeuwen, Moed, Tijssen, Visser, & Van Raan, 2001). Ils montrent que depuis 1980, l'anglais n'a fait que s'imposer davantage, ce qui tend à confirmer les prédictions de Garfield sur l'homogénéisation nécessaire à la circulation des connaissances (Figure 17).

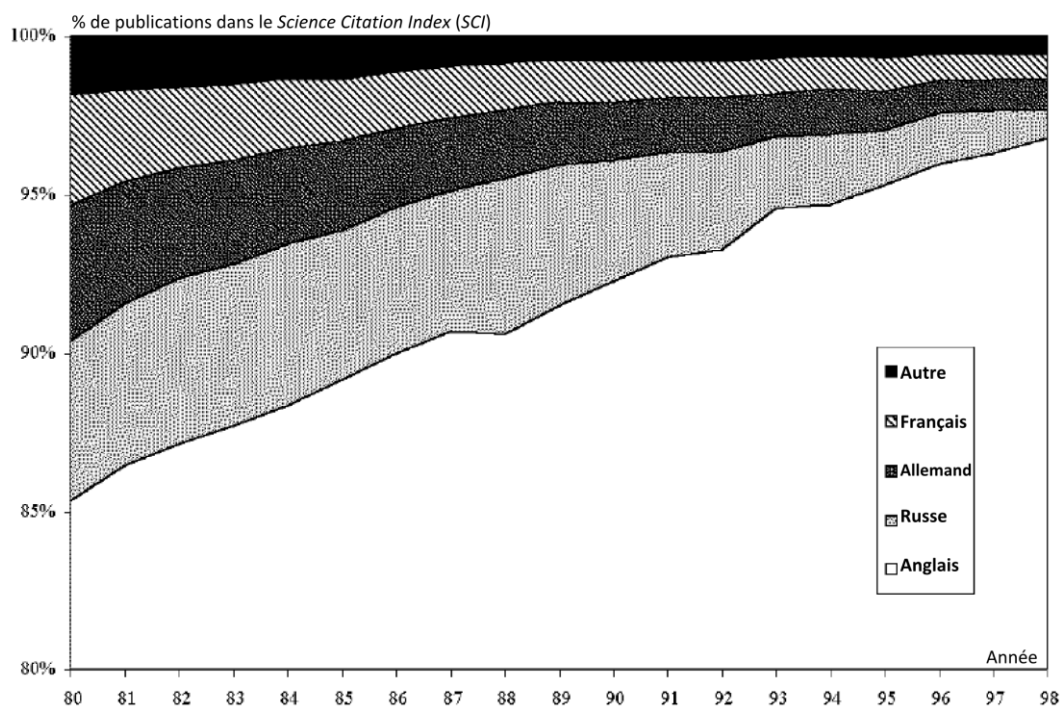


Figure 17 – Évolution des langues dans le SCI, 1980-1998. Figure extraite et traduite de l'article « *Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance* » (van Leeuwen *et al.*, 2001)

Comme le système de référencement conçu par Garfield est fondé sur l'impact des revues (le nombre de citations qu'elles reçoivent) et que les revues en anglais sont les revues qui reçoivent le plus de citations dans le système (n'oublions pas qu'il y a un biais américain dès l'origine), ce sont les revues en anglais qui ont le plus de chances d'être citées et ajoutées au périmètre du SCI. Parallèlement, ce sont les revues qui ne sont pas en anglais qui sont les plus susceptibles d'être exclues du périmètre du SCI puisqu'elles sont

⁴² L'original : « As long as peer-reviewed journals are a good tool for scientific communication, and hence for description of the international scientific landscape in most disciplines, it should be remembered that rather different pictures of world science may be drawn depending on the choice of database and its perimeter of inclusion. » (Zitt *et al.*, 2003).

moins citées. Aussi, van Leeuwen *et al.* démontrent que le fait de comptabiliser les revues non-anglophones qui persistent à exister dans la base pour mesurer l'impact scientifique de la France et de l'Allemagne est en fait défavorable au score respectif de ces deux pays. Ces réalités mathématiques sont entretenues par le biais initial de la base en faveur des États-Unis : dans le monde du *Web of Science*, les productions en anglais sont donc la norme « internationale ».

Finalement, la notion de « science internationale » est une notion rarement explicitée, ce qui contribue à camoufler le fait que par « internationale », il est souvent suggéré « utilisant la langue anglaise qui est la langue internationale ». Or, il est problématique de considérer qu'une revue en anglais est « internationale » par définition. Avant de parler de revues internationales, il faudrait toujours préciser ce que recouvre cette appellation. Les auteurs qui comme Michel Zitt se sont penchés sur la question ont montré que plusieurs facteurs pouvaient être pris en compte : la nationalité des auteurs, la nationalité des citations reçues, la nationalité des abonnés à la revue, la nationalité des auteurs de co-publications, la langue et le pays d'édition, ou encore la nationalité des membres du comité éditorial. Soucieux de tenir compte des déséquilibres initiaux de la base, Zitt a essayé de mesurer l'internationalité des revues à l'aide d'indices normalisés. Pour lui, le degré d'internationalité d'une revue 80 % ukrainienne et 20 % biélorusse ne devrait pas être le même que celui d'une revue 80 % américaine et 20 % britannique (Bassecoulard & Zitt, 1998). Dans son raisonnement qui est cohérent avec une approche fondée sur la fabrique d'indicateurs d'impacts nationaux, la seconde revue est plus internationale parce qu'elle dispose d'un plus large impact. Le problème c'est qu'une simple normalisation conduit à traiter toutes les revues à dominante américaine comme « internationales » parce qu'elles correspondent au profil normal ou moyen dans la base, c'est-à-dire le profil le plus fréquent et par conséquent celui par rapport auquel les autres revues s'écartent plus ou moins. Puisque cet extrême n'est pas non plus acceptable, Zitt cherche une manière de concilier plusieurs indices.

En fait, les scientomètres distinguent un système dominé par la « culture de la citation » (l'expression est de Wouters, 1999), auquel ne participent que ceux qui ont les moyens et la volonté de jouer au jeu dont les règles ont été fixées aux États-Unis, et un système élargi, plus hétérogène, comprenant ceux qui ne « jouent pas dans la même cour ». Ainsi, dans notre revue de littérature sur la question de la nationalité des revues, nous n'avons trouvé qu'une seule étude bibliométrique qui traite sur un pied d'égalité des revues d'Amérique Latine et des revues d'Amérique du Nord. Il s'agit de celle de Gualberto Buena-Casal *et al.* qui revient à la définition littérale d'« international » et défend l'idée que l'anglais n'est pas une condition nécessaire de l'internationalité. Cette étude qualifie d'internationale toute revue qui est le produit de la collaboration d'au moins deux

pays (Buela-Casal, Perakakis, Taylor, & Checa, 2006). Dans l'échantillon de revues analysé, limité à quelques spécialités, les revues en espagnol publiées en Amérique du Sud sont plus souvent internationales que les revues en anglais. Cette réflexion sur la notion de « revue internationale » en psychologie comportementale et en physique théorique rejoint celle engagée par plusieurs géographes au cours des années 2000, en vue de questionner l'hégémonie anglo-américaine en géographie humaine. Parmi eux, Javier Gutiérrez et Pedro López-Nieva (Gutiérrez & López-Nieva, 2001) se demandent si les « revues internationales » de géographie humaine sont réellement internationales (comme l'indique le titre de leur article), ce qui les conduit, en fait, à vérifier si les revues de géographie humaine en anglais sont réellement internationales. Ils réalisent la transition entre les deux questions à l'aide d'une pirouette rhétorique : « la langue de toutes les revues sélectionnées est l'anglais (il est largement accepté que l'anglais est la langue des revues internationales) »⁴³ (*ibid.*). Quelques années plus tard, Maria-Dolors Garcia-Ramon pose le problème de l'internationalité de la discipline et incite les géographes à communiquer et à faire l'effort d'internationaliser les revues, pour éviter aux géographes d'évoluer au sein de périmètres géographiques étriqués (Garcia-Ramon, 2003). En 2006, le problème étant toujours bien présent, le Néerlandais Manuel B. Aalbers et l'Italien Ugo Rossi plaident en faveur du développement d'une géographie européenne pour faire contrepoids à l'hégémonie anglo-américaine (Aalbers & Rossi, 2007).

Très souvent, derrière la question de « l'internationalité » se cache celle de la visibilité et à ce titre, la mesure clef est la mesure d'impact qui dépend du nombre de citations reçues. La difficulté avec laquelle les scientomètres doivent composer vient du fait que l'impact n'est pas indépendant de la géographie puisqu'il varie avec les opportunités de collaboration, la langue de publication ou encore le caractère « *mainstream* » des problèmes de recherche abordés. Ainsi, outre les grandes différences entre domaines disciplinaires (qui justifient la normalisation par domaines pour les mesures d'impact), il a été prouvé que les articles écrits en collaboration sont plus cités que les autres, et d'autant plus s'il s'agit d'une collaboration entre auteurs de pays différents (Narin, Stevens, & Whitlow, 1991 ; Frenken, Hözl, & Vor, 2005). Comme l'explique le scientomètre néerlandais Paul Wouters, auteur d'une thèse sur la « culture de la citation » (*op. cit.*), en concevant un système de référencement fondé sur les liens de citations entre publications, les spécialistes de l'information ont créé une représentation de la science (dans laquelle « toutes les citations se valent »), qui n'est pas sans effet sur les représentations que les scientifiques et les politiciens se font de la science et sur la science elle-même (Wouters, 1998). Ce processus complexe est décrit par Paul Wouters à l'aide du vocabulaire constructiviste de la

⁴³ L'original : « The language of all the journals selected is English (it is widely accepted that English is the common language of international journals). » (Gutiérrez & López-Nieva, 2001).

« traduction » : l'index de citations donne une image systémique de la science à partir d'un ensemble de revues scientifiques qui ne donne, lui-même, qu'un aperçu de la science. Ce processus est un peu comparable à celui bien connu des géographes de mise en carte du monde. La fixation d'un certain nombre de paramètres tels que la position du méridien de référence (le méridien de Greenwich), le choix d'une projection, ou les limites des continents ont un rôle déterminant sur la représentation du monde véhiculée auprès d'un très large public au niveau mondial (Grataloup, 1999 ; Pumain & Saint-Julien, 1997, p. 9). De la même manière, avec le projet de produire des atlas de la science et avec la possibilité de faire varier les périmètres de couverture des index, les scientomètres ont le pouvoir de configurer notre représentation de la science et, au-delà, alors qu'ils ne l'avaient pas vraiment envisagé à l'origine, notre image de la science dans le monde.

Tandis qu'« internationalité » et « visibilité » sont souvent confondues, « visibilité » et « qualité » sont deux notions devenues synonymes dans la « culture de la citation ». En effet, les indicateurs de qualité scientifique et les classements sont nombreux à s'appuyer sur le nombre de citations reçues de la part d'articles indexés dans le WoS. D'après Wouters, l'ambition de cette approche est de déduire une hiérarchie avec objectivité sans s'occuper du contenu proprement dit de la littérature. Wouters explique : « Certes, les scientifiques ont été sommés de rendre compte de leurs recherches depuis longtemps déjà. Mais ces évaluations n'ont jamais dépassé le cadre des frontières disciplinaires. La vraie valeur de la recherche dans un domaine ne pouvait être jugée que par des experts de ce domaine. La qualité de la recherche ne pouvait être envisagée qu'à travers le langage propre au domaine. Cette limite est arrivée à son terme avec l'émergence de la culture des citations. C'est la première fois qu'une représentation de la science émerge dans laquelle un expert extérieur peut juger de la qualité de la science à propos de n'importe quelle spécialité. »⁴⁴ (Wouters, 1998, *op. cit.*). Pourtant, un article écrit à plusieurs mains, en anglais, sur un sujet qui intéresse des scientifiques de plusieurs nationalités a sans doute plus de chance d'être cité alors que ce n'est pas nécessairement un article de qualité, et nous pouvons très bien imaginer qu'un article qui remplisse toutes les conditions inverses le soit. Se fier naïvement aux indicateurs de citation, c'est postuler que la science n'a pas de frontières disciplinaires et géographiques alors que cela reste à prouver.

L'analyse de Wiebke Keim sur la représentativité des sciences sociales issues des pays du Sud dans les bases de données de référence donne à réfléchir en ce sens (Keim,

⁴⁴ L'original : « Certainly, scientists have been required to account for their research for a long time already. But this accounting never surmounted the compartmentalization of science. The real value of research in a certain field could only be judged by the experts in that field. The research quality could only be discussed in the relevant field-specific language. This has come to an end with the emergence of the citation culture in science. It is the first time a representation of science emerged in which an expert outsider's judgment on the quality of science can be formed in any scientific specialty. » (Wouters, 1998)

2010). Au sujet du cas africain, il conclut : « Cette brève analyse indique un manque de consensus par rapport à la sélection de revues africaines qui seraient parmi "les plus importantes du monde" : aucune revue n'est contenue dans les trois bases. Le fait que le JCR (*Journal Citation Report*), aussi bien que FRANCIS, ignorent deux des revues les plus anciennes et les plus reconnues au niveau continental, les publications du CODESRIA *Africa Development* et l'ancien *South African Sociological Review* – aujourd'hui *African Sociological Review* –, devrait discréditer ces bases aux yeux de la communauté scientifique africaine. D'un autre côté, une revue populaire sans comité de lecture, le *South African Labour Bulletin*, apparaît dans l'une des bases de données, ce qui souligne l'ignorance commune du secteur des publications en sciences sociales africaines à Philadelphie, Cambridge et Paris. Les bases de données confirment donc l'hypothèse que les sciences sociales africaines sont hautement marginalisées par rapport au "mainstream" international. » (*ibid.*). Pour Keim, en même temps qu'elles « visibilisent », les bases de données bibliographiques ont le pouvoir de marginaliser. Il considère qu'il faut les appréhender comme « un instrument de marginalisation qui renforce la domination nord-atlantique » (*ibid.*). Il nuance toutefois la portée de ses résultats car il a conscience du manque de données empiriques disponibles pour étayer sa critique : « Néanmoins, l'analyse esquissée jusqu'ici n'exclut pas la possibilité que la faible visibilité de la production des pays du Sud (en excluant, ici, l'Australie), corresponde à un problème réel de sous-développement scientifique, plus particulièrement du secteur de l'édition et qui se répercute sur la quantité de publications ; en ce cas, les résultats seraient, en effet, en corrélation avec la production scientifique. On ne pourra donner qu'une réponse partielle à cette proposition, car il n'existe pas de sources alternatives de données à l'échelle mondiale. » (*ibid.*).

En effet, le problème avec la représentativité c'est qu'il n'existe pas de moyen évident pour la mesurer : il n'y a pas de moyen parfaitement fiable de connaître le nombre de revues qui ne sont pas couvertes, puisque leur particularité est justement de manquer de visibilité, sans compter qu'il n'est pas toujours pertinent de traiter des problèmes de représentativité en prenant le niveau national comme niveau de référence, puisque l'aire d'influence des revues ne coïncide pas forcément avec les cadres nationaux. En dépit de ces difficultés, il existe quelques travaux qui tentent d'estimer le degré de représentativité des données de l'ISI Thomson Reuters. Parmi eux se trouve l'étude sans concessions de Mark P. Carpenter et Francis Narin qui a mis en évidence, dès le début des années 1980, la très bonne représentation des États-Unis et du Royaume-Uni, ainsi qu'une sous-représentativité de l'URSS et du Japon dans le SCI, y compris en biomédecine (Carpenter & Narin, 1981).

Plus près de nous, à la suite de sa contribution au rapport de la *Royal Society* sur le paysage scientifique mondial en 2011, Caroline Wagner s'est intéressée avec Shing Kit

Wong à la sous-représentation de quelques pays en forte croissance scientifique : le Brésil, la Russie, l'Inde et la Chine qu'elle désigne à travers l'acronyme « BRIC » (Wagner & Wong, 2012). Pour chacun des pays considéré dans leur étude, Wagner et Wong établissent que le rapport entre le nombre de revues couvertes par le SCI Expanded et le nombre de revues potentiellement couvertes est proche de 3 %. Or, elles estiment, par ailleurs, que sur l'ensemble des revues mondiales citées dans le *Journal Citation Report* (JCR), seulement 3 % sont recensées dans le SCI Exp, arguant que si les BRICs étaient sous-représentés ce pourcentage devrait être plus faible pour eux. En effet, puisque le SCI Expanded couvre à la fois 3 % des revues mondiales, et 3 % des revues issues des « BRICs », il ressort que les « BRICs » ne sont pas sous-représentés par l'ISI. Malgré l'intérêt de cette analyse et des ordres de grandeur qui en ressortent, on peut objecter que cela ne résout pas la question de la pertinence du contenu scientifique couvert (soulevée par Keim), et d'autre part regretter que le concept de « BRIC » soit mobilisé sans discussion quant au sens de son application à l'étude des activités scientifiques. En effet, repérer des traits communs à la production scientifique des pays de ce groupe présenterait un plus grand intérêt s'il était prouvé (dans le cadre de l'article ou de ses références), que la stratégie internationale mise en œuvre par les BRICs en matière économique s'étend à leur développement en matière scientifique. En effet, il faut rappeler que cette appellation est apparue au début des années 2000 dans le cadre d'un rapport pour la Goldman Sachs sur les marchés émergents avant d'être réappropriée par les dirigeants des pays ainsi désignés, lesquels en ont fait un outil d'action en matière de politique économique au niveau international. Comme l'explique bien Pascal Marchand dans le cadre d'une étude récente sur la stratégie géopolitique de la Russie, ces derniers tiennent un sommet tous les ans depuis 2009, qu'ils ont élargi à l'Afrique du Sud en 2011 pour devenir les « BRICSs » (Marchand, 2014). Or, il est important de préciser que Wagner et Wong ne font pas référence à la dimension politique de cet objet d'analyse, mais traitent de ce groupe comme d'un ensemble de « pays en voie de développement », ce qui est assez discutable étant donné le niveau de « développement » déjà atteint par ces pays, notamment celui de la Russie qui a une trajectoire bien spécifique. Enfin, bien qu'elles insistent sur la bonne représentativité des BRICs dans le SCI Expanded, Wagner et Wong reconnaissent l'existence d'un fossé entre la représentativité et la visibilité de ces pays, le niveau d'impact (en nombre de citations reçues) par les BRICs restant éloigné de leur niveau de production.

Pour réduire l'inégale visibilité des revues selon leur origine, Wagner et Wong préconisent en conclusion le développement de l'*open access* permettant l'accès en ligne aux versions électroniques des revues scientifiques et émettent l'hypothèse que les outils de traduction automatique comme *Google Translate* pourraient faciliter leur circulation. Contrairement à Garfield, elles considèrent qu'il est irréaliste d'imaginer que tous les articles

soient publiés en anglais, surtout les articles ayant une visée locale, d'autant plus qu'ils peuvent bénéficier des méthodes de traduction automatique. En soulignant le potentiel de l'*open access*, Wagner et Wong rejoignent le point de vue d'un spécialiste indien de « la science dans les pays en développement », Subbiah Arunachalam, qui considère que le libre accès à l'information scientifique pourrait être un moyen de combler le fossé de l'information (Arunachalam, 2005). S'ils ont un idéal en commun – Arunachalam reprend à l'OCDE l'expression « village global de la recherche » et parle de « socialisme éclairé » et de « démocratisation du savoir » tandis que Wagner parle d'une participation généralisée au « collège invisible global » (Wagner, 2008) – Arunachalam est moins optimiste que Wagner : le premier craint que la tendance actuelle à la privatisation de l'information vienne compliquer son rêve d'égalité tandis que la seconde se figure plutôt que dans un contexte où les budgets nationaux de la recherche tendent à se compresser, les opportunités de collaborations offertes par l'intégration au « collège invisible » peuvent améliorer l'efficacité du système scientifique dans son ensemble en poussant à la qualité et en évitant les découvertes redondantes. Ainsi, on pourrait décrire les points de vue en présence comme « de gauche » pour le premier et « libéral » pour le second (au sens français du terme impliquant le « laisser faire »).

Même si l'argument suivant lequel il n'est pas question pour l'ISI Thomson Reuters de couvrir l'intégralité des revues scientifiques mondiales soit toujours d'actualité, cela n'empêche pas l'institut de continuer à élargir le périmètre des revues indexées. De fait, confronté à la concurrence de *Scopus* ou encore de *Google Scholar*, les gestionnaires du WoS ont lancé une opération d'élargissement géographique de la couverture des bases entre 2006 et 2008 baptisée « *Regional Expansion* ». Il s'agit officiellement d'intégrer des revues ayant une visée et un impact « international » moindre mais possédant un « intérêt régional » fort (Testa, 2011). Cette opération montre que l'ISI Thomson Reuters fait co-exister plusieurs stratégies qui peuvent entrer en contradiction : avoir la prétention de couvrir la science qui répond aux « standards internationaux » et élargir sa clientèle. Sur les 655 titres qui ont bénéficié de l'opération d'extension géographique, les données transmises par Guillaume Rivalle de Thomson Reuters nous indiquent que 464 ont été ajoutés au SCI Expanded, 187 au SSCI et 63 à l'AHCI. En même temps, la sélection de nouvelles revues et l'exclusion d'anciennes revues se sont poursuivies dans les conditions habituelles, ce qui a contribué à faire croître la couverture d'ensemble du *Web of Science* (Tableau 6). Ce tableau est important pour ce qui va suivre car il donne l'évolution récente de la couverture pour deux ensembles d'années (1999-2001 et 2006-2008). Ce sont les séries temporelles que nous allons considérer pour mesurer l'évolution de la production mondiale contemporaine par la suite.

Nombre de revues indexées	2000*	2007*	Taux d'évolution
<i>Science Citation Index (SCI)</i>	5822	6659	14,38
<i>Social Science Citation Index (SSCI)</i>	1250	1405	12,40
<i>Art & Humanities Citation Index (AHCI)</i>	1052	900	-14,45
<i>Web of Science (WoS)</i>	5994	8133	35,69

*moyenne mobile sur 3 ans

Tableau 6 – Nombre de revues indexées par l'ISI Thomson Reuters en 2000 et 2007 (moyenne mobile sur 3 ans)

Bien que l'extension régionale représente une partie assez marginale de la croissance d'ensemble des index, nous nous sommes intéressée à sa géographie. Sans prétendre échapper à certaines limites évoquées plus haut, nous avons pris le parti d'analyser la distribution par pays de cet ensemble de revues (622 d'entre elles sur 655) à l'aide de deux indicateurs : les pays d'édition et les pays où sont localisés les auteurs ayant contribué aux revues en 2008 (le deuxième indicateur est le résultat, pour une revue donnée, d'un fractionnement par nationalité des contributions à cette revue). Comme en témoigne le Tableau 7, les deux indicateurs coïncident assez bien (coefficient de corrélation = 0,90, $p\text{-value} = 2,2\text{e-}16$). Il faut dire que pour ces revues ayant une forte dimension régionale, les contributions sont assez largement issues du pays d'édition, y compris pour les Pays-Bas qui d'habitude font exception. En effet, parce qu'ils abritent la maison d'édition multinationale Elsevier, les Pays-Bas sont le siège de nombreuses revues qui ne mettent pas forcément les scientifiques néerlandais à contribution. Mais dans le cas de l'extension régionale, il est plus probable que les revues concernées soient issues d'associations scientifiques ou de maisons d'édition universitaires plutôt que d'éditeurs privés comme Elsevier. En effet, comme le remarquent Gutierrez et Lopez Nieva au sujet de la géographie humaine (*op. cit.*), les éditeurs privés ont tendance à se spécialiser dans les contenus « *mainstream* » au lieu des contenus à « forte dimension régionale » qui sont ici directement visés par l'opération d'extension.

Pays	Nombre de revues par pays d'édition	Répartition par lieu d'édition (%)	Répartition d'après la nationalité des contributions* (%)
Australie	49	7,88	6,15
Allemagne	49	7,88	7,13
Espagne	40	6,43	5,78
Pologne	35	5,63	4,91
Brésil	31	4,98	5,03
Chine	29	4,66	4,93
Royaume-Uni	24	3,86	2,85
France	22	3,54	3,89
Italie	22	3,54	3,13
Japon	22	3,54	4,23
Inde	20	3,22	4,65
Corée du Sud	20	3,22	2,67
Roumanie	17	2,73	2,26
Afrique du Sud	17	2,73	2,19
Turquie	17	2,73	3,05
Lituanie	16	2,57	1,96
Mexique	14	2,25	2,03
Chili	11	1,77	1,02
Croatie	11	1,77	1,24
Iran	11	1,77	2,01
Pays-Bas	10	1,61	1,12
Russie	8	1,29	1,96
République Tchèque	7	1,13	1,22
Slovaquie	7	1,13	0,76
Taiwan	7	1,13	1,02
Belgique	6	0,96	0,72
Colombie	6	0,96	0,82
Slovénie	6	0,96	0,86
Suisse	6	0,96	0,72
Autriche	5	0,80	0,86
Bulgarie	5	0,80	0,48
Hongrie	5	0,80	0,91
Etats-Unis	5	0,80	6,06
Canada	4	0,64	1,22
Malaysia	4	0,64	0,41
Nouvelle-Zélande	4	0,64	0,71
Norvège	4	0,64	0,38
Pakistan	4	0,64	0,36
Serbie	4	0,64	0,56
Singapour	4	0,64	0,20
Venezuela	4	0,64	0,58
Estonie	3	0,48	0,41
Finlande	3	0,48	0,44
Philippine	3	0,48	0,16
Argentine	2	0,32	0,63
Danemark	2	0,32	0,44
Grèce	2	0,32	0,44
Israël	2	0,32	0,48
Lettonie	2	0,32	0,31
Bahrain	1	0,16	0,05
Bangladesh	1	0,16	0,15
Bosnie Herzégovine	1	0,16	0,13
Bielorussie	1	0,16	0,07
Islande	1	0,16	0,13
Macédoine	1	0,16	0,12
Nigeria	1	0,16	0,38
Portugal	1	0,16	0,30
Arabie Saoudite	1	0,16	0,12
Thaïlande	1	0,16	0,31
Ukraine	1	0,16	0,38
Total	622	100	100

*Calculée à partir de la somme fractionnée de participation aux 622 revues

Tableau 7 – L'opération « *Regional Expansion 2006-2008* ». Source : ISI Thomson Reuters

Trois remarques peuvent être faites à partir du Tableau 7 :

- premièrement, les pays étudiés par Caroline Wagner sont parmi les premiers concernés par l'extension régionale, ce qui suggère qu'aides en cela par une logique commerciale, l'ISI a bien l'intention de suivre les dynamiques globales en matière de science, et c'est ce qui en fait un instrument intéressant pour étudier l'évolution du paysage scientifique mondial ;

- deuxièmement, l'Australie est le premier bénéficiaire de cette opération avec l'Allemagne. Comme l'Australie est un pays très dépendant des États-Unis en matière éditoriale, il se peut bien que sa part de production locale ait été sous-représentée avant cette opération (Keim, *op. cit.*). De son côté, l'Allemagne est une force éditoriale importante abritant la maison d'édition privée Springer, mais il est difficile de savoir dans quelle mesure ce facteur joue ici un rôle ;

- troisièmement, comme il était prévisible, le nombre de revues éditées aux États-Unis à avoir profité de l'extension est faible (0,80 %). Cependant, la part des contributions états-uniennes dans les revues bénéficiaires (à fort « contenu régional ») n'est pas négligeable (6 %).

Le choix de l'ISI Thomson Reuters de distinguer l'opération d'extension régionale du processus normal d'élargissement de la base d'après les critères habituels est tout à fait légitime compte tenu des réserves de Michel Zitt concernant l'intégration des revues à orientation nationale pour interpréter l'évolution des indicateurs de productivité nationales. Cette distinction permet aux scientomètres de rendre leurs mesures plus consistantes puisqu'elle leur donne la possibilité d'exclure l'extension régionale de leurs calculs et ainsi de travailler à partir de périmètres aux contenus plus homogènes.

À notre sens, s'il est louable, l'idéal du scientomètre consistant à saisir les réalités scientifiques toutes choses égales par ailleurs est tout aussi difficile à atteindre que celui de l'économètre qui entend construire des modèles de comportement dans l'espace qui soient valables en tous temps et en tous lieux. Ces intentions ont une valeur scientifique indéniable mais ce n'est pas le type d'approche qui motive cette thèse. Puisque notre intention est de capturer le paysage scientifique de l'espace mondial, il est peu pertinent de chercher à annuler les effets liés à l'hétérogénéité culturelle et géographique en travaillant à partir d'un ensemble de revues au profil homogène. Au contraire, en regardant de plus près l'évolution du périmètre du WoS, il nous est apparu qu'il avait un sens au moins pour estimer l'état des efforts nationaux.

En comparant la répartition par pays des publications en 2008 sur la base du périmètre du WoS en 2000 et sur la base du périmètre du WoS en 2008, nous avons constaté que le nouveau périmètre bénéficiait aux pays émergents (même indépendamment de l'extension régionale) ce qui laisse penser qu'il n'y a pas besoin d'extension régionale pour

que l'évolution « naturelle » de la couverture du WoS suive l'évolution réelle des dynamiques scientifiques mondiales (Tableau 8).

Publications du WoS (comptages fractionnés)	A	B	C	Différence
	Répartition des publications en 2000* (%)	Répartition des publications en 2007* (%)		C-B
	Périmètre du WoS en 2000*		Périmètre du WoS en 2007*	
Etats-Unis	30,04	26,22	25,64	-0,59
Chine	2,76	7,83	7,47	-0,36
Japon	7,94	6,28	6,13	-0,15
Royaume-Uni	7,62	5,89	5,91	0,02
Allemagne	6,81	5,65	5,58	-0,06
France	4,95	4,11	4,05	-0,05
Canada	3,46	3,48	3,47	-0,01
Italie	3,22	3,31	3,36	0,05
Espagne	2,28	2,60	2,60	0,00
Inde	1,87	2,44	2,51	0,07
Corée du Sud	1,38	2,34	2,35	0,00
Australie	2,13	2,14	2,19	0,05
Russie	2,64	1,24	1,83	0,59
Brésil	1,07	1,51	1,79	0,27
Pays-Bas	1,84	1,18	1,72	0,54
Taiwan	1,10	1,72	1,69	-0,04
Turquie	0,65	1,36	1,41	0,05
Pologne	0,97	0,87	1,22	0,35
Suède	1,43	1,15	1,12	-0,02
Suisse	1,21	1,11	1,09	-0,01
Belgique	0,91	0,91	0,90	-0,01
Israël	0,95	0,81	0,80	-0,01
Iran	0,14	0,64	0,73	0,09
Grèce	0,49	0,70	0,71	0,01
<i>Nombre de publications</i>	<i>830863,00</i>	<i>1013401,67</i>	<i>1205164,67</i>	<i>191763,00</i>

*Moyenne mobile sur 3 ans. Source: WoS, tous types de documents

Tableau 8 – L'effet du changement de périmètre du Web of Science entre 2000 et 2007 sur la répartition par pays des publications scientifiques totales en 2007 (part mondiale, moyenne mobile sur 3 ans)

Au-delà de l'aspect géographique, il existe aussi, comme l'ont montré Peder Olesen Larsen et Markus von Ins, une raison scientifique justifiant l'expansion de la couverture et la prise en compte de cette expansion dans l'analyse : « Parce que les nouvelles communautés de recherche lancent souvent de nouvelles revues pour diffuser leurs résultats, un ensemble fixe de revues sous-estime les recherches qui ne sont pas encore bien établies. Plus la période étudiée est longue, moins les raisonnements à périmètre constant de revues sont représentatifs de l'évolution de la production scientifique mondiale »⁴⁵

⁴⁵ L'original : « Because new research communities often spawn new journals to disseminate their research findings, a fixed journal set under-represents the type of research that were not already well estab-

(Larsen & von Ins, 2010). D'ailleurs, ces mêmes auteurs attirent l'attention sur le fait qu'en s'élargissant moins vite que celles des concurrents, la couverture du SCI tend à passer à côté d'une partie importante de la production en informatique et en sciences de l'ingénieur. De plus, les chercheurs ont de plus en plus volontiers recours à des modes alternatifs de diffusion de leurs résultats (l'*open archives*, les pages internet personnelles, les réseaux sociaux académiques), ce qui pose la question de l'avenir d'une source comme le SCI pour l'établissement d'indicateurs de performance scientifique (*ibid.*).

Cela étant, la croissance des participations chinoises nous informe sur l'activité de ce pays à la fois du point de vue de son effort scientifique mais aussi vis-à-vis de son ambition de « visibilité ». Aussi, le rapport d'Éric Archambault, directeur de la petite société canadienne Science-Metrix, montre qu'il est possible de capturer les effets de plusieurs événements géopolitiques survenus à partir des années 1980 à partir des données du WoS (Archambault, 2010). Même Michel Zitt et ses collègues, conscients des limites de la base pour mesurer des indicateurs d'impact, ont montré le potentiel de celle-ci pour analyser l'évolution de la participation de grandes puissances scientifiques à l'entreprise mondiale (Zitt, Bassecoulard, & Okubo, 2000). La base est évidemment biaisée, mais ces biais sont significatifs de l'état des rapports de force mondiaux et c'est ce qui en fait un laboratoire intéressant pour étudier le paysage de la science dans le monde.

Même l'absence de participation peut être significative. Elle peut s'interpréter de plusieurs façons si l'on suit le raisonnement de Keim (*op. cit.*) puisqu'il peut s'agir d'une réelle incapacité à peser dans la production scientifique mondiale ou bien d'une stratégie délibérée de ne pas contribuer au système tel qu'il est orchestré par les États-Unis. À l'appui de cette dernière stratégie en sciences sociales, Keim a forgé le concept de « courant contre-hégémonique » qu'il définit de la façon suivante : « Le concept (...) est basé sur l'idée que la domination nord-atlantique dans les sciences sociales repose sur l'acceptation partagée d'une arène de compétition commune. C'est seulement dans la mesure où les institutions et les médias de communication du "mainstream" dominant sont reconnus comme l'arène de compétition centrale de la discipline que la lutte pour la reconnaissance internationale peut commencer et que la marginalisation peut opérer. Par conséquent, la seule manière de développer un "potentiel contre hégémonique" réel et pratique – en opposition aux critiques théoriques présentées plus haut – consisterait à refuser la participation à cette arène commune et le refus de nommer cette arène comme l'arène centrale d'une discipline. Cela arrive, plus souvent dans la pratique même de la recherche et de l'enseignement que dans la discussion théorique : dès qu'une communauté scientifique suffisamment grande tourne le dos à la soi-disant "communauté internatio-

lished at the outset of the period. The longer the period being studied, the less adequate a fixed journal set becomes as a representation of the world's articles throughout the period. » (Larsen & von Ins, 2010).

nale" pour s'orienter vers des arènes alternatives. » (*ibid.*). Si elle peut étonner, cette stratégie n'est pas sans pertinence puisque certains chercheurs spécialistes des pays « non-hégémoniques » ont montré que participer à tout prix à la science « *mainstream* » pouvait être nuisible à l'intégration nationale et à la construction d'une « communauté scientifique nationale ». C'est du constat d'une déconnexion entre les intérêts scientifiques localisés et les intérêts des groupements transnationaux que vient l'expression : « *publish globally, perish locally* » (Hanafi, 2011). Ces dernières considérations montrent bien, s'il le fallait encore que la géographie tient une place fondamentale dans la pratique des activités scientifiques.

En résumé, de nos jours, être indexé dans le WoS, compte tenu du très faible nombre de revues non anglophones qui y sont incluses, suppose de publier en anglais. Les chercheurs chinois l'ont bien compris qui ont produit presque autant de publications recensées dans la base que les Américains au cours des cinq dernières années⁴⁶ (Zhou & Leydesdorff, 2007). Puisque le nombre de revues chinoises en chinois indexées dans la base est très faible (en dépit de l'extension régionale, voir Wu, Fu, & Zhang, 2008), cela n'a été possible que parce que les chercheurs chinois se sont mis à publier dans des revues étrangères en anglais, ce qui a été facilité par leur intégration au sein de comités éditoriaux de revues américaines (García-Carpintero, Granadino, & Plaza, 2010). Aussi, comme l'a démontré Arpana Basu, la densité d'articles par revue varie selon les pays si bien que l'exceptionnelle croissance chinoise de ces dernières années a également été favorisée par une augmentation de la part moyenne d'articles publiés par revues (Basu, 2010). En s'appuyant sur l'exemple de l'Inde et de Vienne en Autriche, Basu montre néanmoins, en accord avec nos propres résultats, qu'il existe une relation forte entre le nombre de revues couvertes et le nombre de publications par pays, voire par ville (*ibid.*). Il est donc bon de connaître le nombre de revues d'un pays recensées par la base de données (même s'il n'existe pas de définition claire pour établir ce nombre avec certitude) pour éviter de méinterpréter l'évolution du nombre de publications scientifiques de ce pays, ou d'une ville de ce pays.

Tout cela illustre bien le fait que le WoS ne permet de représenter qu'une fraction de la production scientifique mondiale : on parle parfois pour l'évoquer de la science « visible », « *mainstream* » ou encore « mondiale ». Ce dernier qualificatif est utilisé par opposition à une science qui serait « locale ». Comme on vient de le voir, il ne faut pas s'y méprendre, la distinction entre une « science mondiale », à visée prétendument

⁴⁶ Dans le cadre du Festival International de Géographie de Saint-Dié portant sur le thème de la Chine en 2013, nous avons consacré un poster à cette question : Maisonobe, M., Baron, M., & Jégou, L. (2013, octobre). La Chine, nouveau géant de la science mondiale. Poster pour le concours du Festival International de la Géographie de Saint-Dié des Vosges, Saint-Dié-des-Vosges: LISST-CIEU. Consulté à l'adresse <http://archivesic.ccsd.cnrs.fr/GIP-BE/hal-01157945v1>

universelle et une « science locale » destinée à un public plus restreint est largement exagérée. Il y a beaucoup de revues indexées dans le WoS qui, alors même qu'elles sont en anglais, pourraient être qualifiées de « locales » compte tenu de leur thématique et du public qu'elles touchent. Pour cette raison, nous préférons le qualificatif « visible » pour évoquer le « monde scientifique » que ces données donnent à voir. Pour synthétiser, nous envisageons le WoS comme un miroir du « monde scientifique » sous influence américaine, dans la mesure où l'on y repère les parutions des chercheurs qui publient en anglais dans des revues dont la ligne éditoriale répond aux critères de l'ISI, un institut privé dont le siège est à Philadelphie. À l'appui de cette approche, on peut citer Xavier Polanco : « L'utilisation d'une base de données comme l'ISI ne permet de mesurer que la position relative des pays dans la littérature *mainstream* et donc uniquement la part de la science des pays en développement qui est citée et utilisée par les auteurs de cette littérature *mainstream*, qui constitueraient ce que l'on appelle la « communauté scientifique internationale. » (Polanco, 1990, p. 45). Dans le cadre de cette thèse et plus précisément de la deuxième partie, nous admettons que le WoS permet de visualiser un « monde scientifique » parmi d'autres, et c'est celui qu'il s'agit d'explorer. Néanmoins, la méthode de traitement spatial des données bibliographiques définie à partir des données du WoS et avec laquelle nous l'explorons est conçue pour pouvoir ultérieurement s'appliquer à n'importe quel corpus de données bibliographiques, quelle que soit sa source.

PARTIE 2.
LA GÉOGRAPHIE DES SCIENCES À
TRAVERS LE WEB OF SCIENCE

Après avoir démontré dans la partie précédente l'intérêt pour les géographes d'une aventure sur le continent de la science et après avoir envisagé le *Web of Science* (WoS) comme un observatoire digne d'intérêt pour explorer la face visible de ce continent, il est temps de se pencher sur l'aspect technique de l'expédition : comment utiliser le WoS pour appréhender la géographie des activités scientifiques ? Pour plus de clarté, nous reprendrons certains éléments, en particulier sur le contenu et les contraintes de la source, qui ont été évoqués à la marge lors de la précédente partie. Comme la publication est un moyen privilégié de diffusion des recherches et que les collaborations scientifiques donnent de plus en plus souvent lieu à des publications co-signées par plusieurs partenaires, nous considérons que l'exploitation géographique de données bibliographiques est pertinente pour rendre compte à la fois des logiques de production et d'organisation spatiale de la science et de leurs évolutions à l'échelle mondiale. Nous décrivons dans ce qui va suivre la méthode établie en équipe pour passer des données brutes à leur mise en forme avant de procéder à la représentation et à l'analyse. Cette méthode est conçue pour être reproductible et opérationnelle pour traiter tous types de corpus bibliographiques. À plusieurs étapes de sa conception, des choix ont été effectués en équipe, et justifiés à la lueur des objectifs de la recherche. Ces choix, susceptibles d'influencer les résultats, doivent être faits en connaissance de cause. C'est tout l'objet des trois prochains chapitres méthodologiques que de sensibiliser aux enjeux propres à l'analyse spatiale d'une part et à la scientométrie d'autre part.

Ainsi, le travail de délimitation d'entités locales comparables présenté dans le premier chapitre de cette partie (Chapitre 6) repose sur la conviction qu'une fois en possession d'informations sur la répartition géographique des publications et l'organisation spatiale de la production à plusieurs niveaux spatiaux (Chapitre 7), nous pouvons saisir des logiques spatiales propres au fait scientifique et leurs évolutions (Chapitre 8), à commencer par les tendances lourdes en matière de répartition et d'organisation des activités scientifiques dans le monde (Chapitre 9). Une fois que la méthode a été établie et éprouvée pour la totalité des enregistrements de la base sur plusieurs années, il devient possible de la décliner à différentes disciplines ou même à des sous-ensembles de publications ayant une thématique commune, un auteur en commun, ou une référence en commun. Cette méthode de scientométrie spatiale a vocation à s'adapter à un grand nombre de questions de recherche et même à s'appliquer à des données bibliographiques issues d'autres sources que le WoS. C'est la raison pour laquelle le développement méthodologique n'est pas borné par une question ou des hypothèses précises concernant la géographie des sciences. En revanche, son ambition est de proposer une démarche standardisée pour traiter les adresses institutionnelles associées à des publications scientifiques à plusieurs niveaux géographiques.

CHAPITRE 6. LOCALISER LES DONNÉES BIBLIOGRAPHIQUES

À partir de l'information géographique (les adresses ou affiliations institutionnelles) associée à des corpus de publications scientifiques, nous avons cherché à établir une procédure permettant de produire et d'examiner (Eckert *et al.*, 2013) :

- Une liste de lieux ;
- Un nombre de publications scientifiques par lieu et par année ;
- Un nombre de relations entre lieux compte tenu des co-publications par année ;
- Un nombre de relations entre lieux indiqué par les citations par année.

Le défi à relever avant d'y arriver était double. Comment et à quelle résolution localiser les publications et analyser leur répartition (l'unité de recherche, le laboratoire, l'établissement, la municipalité, l'agglomération...) ? Comment agréger l'information brute pour obtenir un nombre de publications par lieu et mesurer une intensité relationnelle à partir des co-publications entre lieux ?

La réponse apportée à la première question traduit la volonté de traiter l'information bibliographique à un niveau de raffinement jamais égalé pour une analyse de la répartition des publications dans l'ensemble du monde : le niveau local (Section 1). Dans cette perspective, nous sommes confrontés au problème, connu en géographie, de la délimitation d'entités spatiales homogènes à l'échelle du monde (Section 2). Pour la seconde question, abordée dans le chapitre suivant (Chapitre 7), il s'agit d'arbitrer entre plusieurs méthodes de comptage et de simplification des données et donc de faire face à un problème classique en scientométrie. Dans l'ensemble, les choix qui sont défendus dans ces deux chapitres témoignent de l'intention de délaisser l'écologie des lieux (ce qui s'y passe, la dynamique et l'organisation interne) pour se concentrer uniquement sur leur situation (leur rang ou place par rapport aux autres lieux) et les relations établies avec les autres lieux. Au fur et à mesure de la présentation des choix méthodologiques, il apparaîtra que d'autres approches sont possibles et que les résultats obtenus sont sensibles aux méthodes d'agrégation et de synthèse des données brutes.

En fait, l'entrée par les données bibliographiques nous invite à considérer tous les lieux désignés par les auteurs de publication comme des lieux de science. D'ailleurs, en décomposant les adresses des auteurs, il est possible d'identifier avec précision ces lieux depuis le niveau de l'établissement jusqu'au niveau national.

1. Les lieux de la production scientifique

« En vérité, il n'y a pas de centre du Monde ni de hiérarchie des lieux mondialisés : la centralité se déplace. C'est le propre de la mondialisation. » (Retraillé, 2012, p. 12)

1.1. Les traitements par organisation

Étudier la répartition géographique de l'activité scientifique demande de s'intéresser aux lieux ou sites de l'activité. À cet égard, les publications ne sont qu'une source parmi d'autres, qui permet de repérer ces lieux à travers les adresses ou affiliations spécifiées par les auteurs. Les contributions à des revues scientifiques sont d'origines variées : elles peuvent aussi bien provenir d'établissements publics ou privés, d'auteurs statutaires ou non statutaires. Le fait de publier peut être vu comme un marqueur, voire un prérequis, du métier de scientifique. Plutôt que d'adopter une définition *a priori* des scientifiques, ce qui est difficile pour une analyse à l'échelle mondiale, l'approche bibliométrique suppose qu'un scientifique est un auteur de publications scientifiques. Ce prémisses est conforté par le travail du sociologue des sciences mertonien Warren Hagstrom qui considère que l'appartenance à « la communauté scientifique » est actée par le fait de publier (Hagstrom, 1965). De manière analogue, Bruno Latour et Steve Woolgar remarquent que : « la production d'articles est le but essentiel » de l'activité (Latour & Woolgar, 1979). Dès lors, que vous soyez affilié à un parc naturel, une université, un observatoire, un hôpital, vous êtes des scientifiques du moment que vous publiez.

Suivant le statut et le lieu de travail des auteurs, l'injonction à publier est plus ou moins marquée. Par exemple, le RIVM (*Netherlands National Institute for Public Health and the Environment*), une organisation néerlandaise sous la tutelle du ministère de la santé, n'oblige pas ses employés à publier ou enseigner. Mais, pour Harry van Steeg, spécialiste de la réparation de l'ADN, membre du RIVM (que nous avons rencontré dans son laboratoire à Bilthoven), le fait de publier et d'enseigner permet d'interagir et de bénéficier d'une légitimité parmi les chercheurs néerlandais de son champ d'expertise. Du reste, la signature d'Harry apparaît dans certains articles de recherche auquel il a peu participé. Parfois, sa seule contribution a été d'envoyer le modèle de souris mutante, appelé « XPA », qu'il a mis au point dans le cadre de ses expériences sur le cancer. Interrogé sur ses collaborations avec une équipe de recherche en Alabama, il explique : « Le modèle XPA est tellement célèbre que j'ai dû l'envoyer partout dans le monde. Certains chercheurs sont si gentils qu'ils mettent mon nom sur leurs articles en guise de reconnaissance.

C'est le cas pour cette équipe d'Alabama que je ne connais pas. »⁴⁷. De ce fait, les adresses associées à chaque publication permettent de repérer les lieux ayant participé formellement à la publication (la géographie institutionnelle), qui ne sont pas forcément ceux où la recherche a été effectuée.

La plupart du temps, les auteurs vont jusqu'à préciser, dans leur adresse, le nom du département universitaire ou de l'unité de recherche à laquelle ils appartiennent. Il est donc concevable d'utiliser des données bibliométriques pour étudier la structuration de l'activité scientifique y compris à l'intérieur d'un équipement précis. Par exemple, nous pouvons étudier la répartition des publications entre équipes de recherche, laboratoires ou départements au sein d'une seule et même université. Si l'exploitation des données bibliographiques à ce niveau de précision n'est pas sans pertinence, elle pose des problèmes liés au manque d'homogénéité des adresses institutionnelles. L'absence d'homogénéité est double : d'une part, il existe de nombreux moyens de désigner un seul et même lieu et d'autre part, la définition, la nature et la taille des unités de recherche varient dans l'espace et le temps.

Le premier problème vient du fait que malgré la multiplication des tentatives de normalisation, l'adresse est déclarée librement par le chercheur. Même si chaque sous-ensemble ou élément de l'adresse est séparé par une virgule, les auteurs ont au moins deux manières de procéder : la hiérarchie suivie par l'écriture de l'adresse peut être ascendante (« sous-organisation, organisation, ville, province, pays ») ou descendante (« organisation principale, sous organisation, ville, province, pays ») (OST, 2007 ; Aubert, 2008, p. 16-18). Outre que la gradation retenue dépend de la libre appréciation du chercheur ou varie selon les cultures disciplinaires et nationales, les dénominations des laboratoires et des établissements sont rarement figées et unifiées. Dès lors, pour étudier la production scientifique au niveau le plus précis, il faut se soumettre à un travail minutieux d'homogénéisation et de désambiguïsation de tous les éléments de texte de l'adresse. Ce travail devient très lourd à mettre en œuvre à partir d'une certaine quantité d'adresses à traiter, c'est pourquoi, nous avons dû réfléchir à des moyens de l'automatiser, on y reviendra.

Le deuxième problème, c'est-à-dire le fait que la taille et la nature des organisations varient d'un endroit à l'autre et d'une génération à l'autre (la taille et l'organisation des laboratoires évoluent dans le temps), se pose surtout dans le cas où l'on voudrait comparer l'activité de plusieurs organisations en même temps ou bien la dynamique de production scientifique interne à une seule et même organisation à plusieurs dates. La compa-

⁴⁷ Verbatim: « XPA model is so popular than I have to send it all over the world. Some scientists are so nice, they put my name on the paper (ex: Alabama). ». Entretien du mercredi 13 juillet 2011 avec Harry van Steeg au RIVM à Bilthoven.

raison dans le temps et l'espace est très délicate entre des entités aussi mouvantes et hétérogènes que les cellules ou unités de travail scientifique.

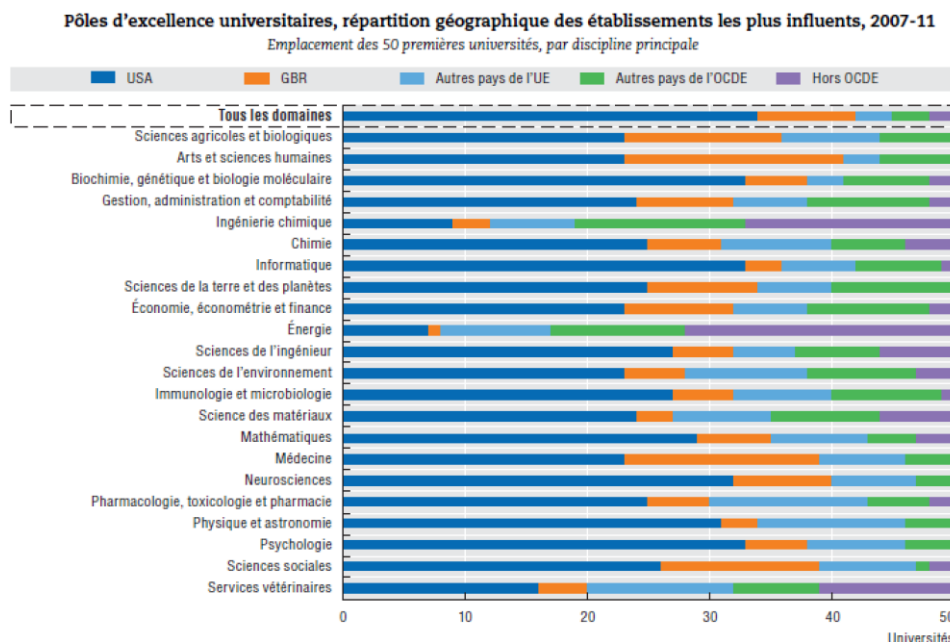
Enfin, même si l'on parvient à localiser l'activité scientifique à un très haut degré de précision en décomposant et en « nettoyant » les adresses associées aux publications scientifiques, la significativité des résultats dépendra au bout du compte de la source mobilisée pour construire le corpus de publications. Au cas où l'on souhaiterait étudier l'activité de production d'un établissement en particulier, la liste de publications issue du WoS est sûrement moins adaptée et complète que la liste que l'on pourrait recueillir en formulant une demande expresse auprès des responsables de l'établissement étudié. En France par exemple, les laboratoires de recherche produisent des rapports d'activité et les chercheurs sont incités à signaler l'ensemble de leurs travaux sur des archives ouvertes, de sorte qu'il serait possible de se procurer une liste de publications plus exhaustive dans la perspective d'une monographie de laboratoire. De la même manière, pour analyser la production scientifique d'un pays ou d'une région, il peut être judicieux de s'adresser à des organismes de recherche, d'évaluation, de documentation ou de statistique plus locaux. S'ils ne disposent pas forcément de données formatées pour être exploitées, certains services possèdent sans doute des données plus complètes que celles du WoS, au moins en ce qui concerne la recherche publique. Qu'il existe d'autres sources d'information ne veut pas dire qu'il soit préférable de les mobiliser à tous points de vue. C'est ainsi qu'en France, malgré les alternatives existantes, l'OST Paris exploite les données payantes du WoS pour calculer des indicateurs de production et situer la place de la science française dans le contexte international. Pour la comparaison ou *benchmarking* à l'échelle de plusieurs pays, il semble effectivement préférable d'avoir recours à une source dont les éléments sont sélectionnés à partir d'un critère constant (pour le WoS : le nombre de citations reçues par les revues recensées) (voir Partie 1. Chapitre 5.2.2). La base *Scopus* présente le même avantage, de sorte que pour comparer le niveau de production scientifique des organisations de recherche et des territoires par-delà les frontières nationales, *Scopus* et le WoS sont en situation de concurrence.

C'est ainsi que pour se doter d'indicateurs synthétiques permettant la comparaison internationales, l'OCDE s'appuie sur les mesures réalisées par SCImago, un groupe de recherche espagnol, à partir des données de la base *Scopus* (OCDE, 2013) ; tandis que la direction recherche de la Commission Européenne se fournit auprès de l'entreprise canadienne Science-Metrix, qui exploite les données bibliographiques de plusieurs bases : le WoS, *Scopus* et la base de publications médicales *Medline* (Campbell, Lefebvre, *et al.*, 2013 ; Labrosse *et al.*, 2013). Les niveaux d'analyse retenus par ces institutions internationales pour faire du *benchmarking* sont celui des organisations mais également ceux des pays et provinces — abordés dans la section suivante.

Pour la comparaison d'organisations, l'analyse porte toujours sur un nombre limité d'organisations. En effet, il serait démesuré de considérer simultanément toutes les organisations car, on l'a vu, même si le remplissage des adresses était normalisé – ce qui faciliterait leur identification – il faudrait composer avec une population d'organisations extrêmement vaste et elle-même très hétérogène, c'est-à-dire avec des entités de tailles et de fonctions variables. Pour dépasser cette difficulté, les analyses bibliométriques multinationales tendent à se centrer sur une seule catégorie d'organisations ou bien à ne traiter que des organisations les plus dynamiques en matière de production. Par exemple, pour le compte de la direction recherche de la Commission Européenne, Science-Metrix a réalisé une étude bibliométrique sur la production scientifique des 100 entreprises privées, multinationales pour beaucoup d'entre elles, qui ont le plus publié dans l'Espace Européen de la Recherche (EER) entre 2007 et 2011 (Campbell, Ventimiglia, & Archambault, 2013). Ce travail est le résultat d'une opération de récupération puis de « nettoyage » des adresses des entreprises et de leurs sous-traitants indexées dans *Scopus*. Les données de publication ont été agrégées et ventilées par entreprises et par secteurs industriels. Dans le cadre du même partenariat entre Science-Metrix et la Commission Européenne, des traitements ont été réalisés au niveau des universités (Labrosse *et al.*, *op. cit.*). Pour ce dernier travail, l'analyse s'est concentrée sur 300 universités appartenant à l'EER sélectionnées d'après leur pays d'appartenance et leur niveau de production. De manière analogue, pour la récente étude de l'OCDE sur la science et l'innovation, le groupe SCImago a proposé une analyse de la répartition par pays et secteurs scientifiques des 50 premières universités en matière d'impact scientifique mesuré et normalisé à partir des données de citations extraites de *Scopus* (OCDE, *op. cit.*). La Figure 18 en est extraite.

Pour ventiler et comparer des données bibliographiques par organisation à l'échelle de plusieurs pays, il est donc fréquent de sélectionner un nombre limité d'établissements, généralement les plus actifs et de ne pas s'encombrer de l'activité de production des petits établissements difficilement repérables dans les adresses institutionnelles. C'est aussi ce qui se pratique dans le cas des classements. La plupart des classements mondiaux d'organisations scientifiques sont des classements d'universités limités aux universités les plus publiantes du monde. Le classement de Shanghai et celui de Leyde sont dans ce cas-là et s'appuient tous deux sur les données du WoS. Ne considérant que la production des universités, ils ont pour ambition de classer les universités suivant leur plus ou moins grande efficacité. Ils reposent sur l'idée, régulièrement attaquée par les adversaires des classements, que les universités sont une population homogène à l'échelle mondiale. Or, l'appellation même d'« université » qui s'est largement diffusée dans la deuxième moitié du XX^e siècle désigne des établissements très différents et sans commune mesure d'un pays à l'autre, voire au sein d'un seul et même pays. De plus, la popularité

de ce type de classements tend à faire oublier que les universités ne sont pas les seules organisations où se pratique une activité scientifique.



Source : OCDE et SCImago Research Group (CSIC), *Compendium of Bibliometric Science Indicators 2014*, d'après Scopus Custom Data, Elsevier, mai 2013. Davantage de données via StatLink.

StatLink <http://dx.doi.org/10.1787/888932931468>

Comment lire ce graphique

L'axe des abscisses indique la répartition géographique des 50 premières universités dans les principales disciplines (axe des ordonnées) selon leur impact normalisé. Le nombre minimal de publications fixé pour les établissements est de 100 documents en 2011, sauf pour certaines disciplines, pour lesquelles il a été fixé à 50 documents. L'influence normalisée correspond au rapport entre le nombre moyen de citations reçues par une unité donnée et la moyenne mondiale des citations au cours de la même période, par type de document et par discipline ; autrement dit, la normalisation se fait au niveau de l'article. Si un article relève de plusieurs disciplines, on calcule une valeur moyenne pour les disciplines en question. L'influence normalisée des établissements est calculée pour la période 2007-11 et ne tient compte que de la production à laquelle le pays concerné apporte la principale contribution (production dont l'auteur principal est rattaché à l'établissement).

Figure 18 – Pôle d'excellence universitaire, répartition géographique des établissements les plus influents. Figure extraite du rapport de l'OCDE « Science, Technologie et Industrie : Tableau de bord de l'OCDE 2013 » (OCDE, 2013, p. 54)

En somme, il est préférable de recourir aux grandes bases de données bibliographiques comme le WoS ou *Scopus* pour des analyses comparatives de la production scientifique à l'échelle mondiale plutôt que pour des monographies d'établissements ou pour des études restreintes à un périmètre géographique limité. Lorsqu'elle est mobilisée pour des classements ou comme outil d'administration scientifique, l'information géographique de ces bases est regroupée soit au niveau des organisations, soit au niveau des pays ou provinces. Le nombre, la variété et l'hétérogénéité des organisations limitent leur comparabilité, mais, bien que les pays et provinces soient des entités plus pérennes que les organisations, leur étendue présente aussi une forte hétérogénéité, c'est l'objet de la prochaine section.

1.2. Les traitements par entité spatiale

À partir des adresses institutionnelles, en dehors des traitements par organisations, il est possible d'agréger l'information au niveau d'entités spatiales. Avec le niveau des organisations qui est, au moins depuis les années 1980, un niveau d'analyse privilégié du point de vue de l'évaluation scientifique et du *benchmarking*, celui des pays est le niveau d'analyse traditionnellement le plus utilisé par les bibliomètres.

Il faut dire que l'OCDE encourage le développement de politiques scientifiques au niveau des États, ce qui a pour effet de donner au pays un caractère structurant pour la pratique et l'organisation des activités scientifiques (Henriques & Larédo, 2013). En particulier, la plupart des États participent au financement de la recherche par le paiement de salaires mais aussi par l'allocation d'investissements ciblés, d'où l'intérêt de mesurer, à travers le volume de publications, le devenir des financements qu'ils consentent. Prendre ce niveau d'analyse pour étudier la répartition des activités scientifiques se justifie à la fois par l'appartenance institutionnelle des scientifiques (Storer, 1970), et par la plus grande simplicité technique de ce type de traitement (Frenken & Hoekman, 2014). En effet, le pays est un champ bien renseigné et normalisé par les gestionnaires des bases de données bibliographiques. Ainsi, les administrateurs du WoS exploitent cette information pour commercialiser des indicateurs et proposer des options d'analyse aux utilisateurs de la base de données en ligne. Voilà donc plusieurs générations que les bibliomètres, particulièrement ceux qui sont habitués à exploiter les données du WoS, traitent de la production scientifique par pays (voir Partie 1. Chapitre 5.2.2). Pourtant, la résolution de ces analyses n'est pas toujours optimale au vu des possibilités de traitement envisageables et de l'ambition des recherches ; d'autant que, malgré leur proximité sur le plan juridique (la plupart disposent d'une reconnaissance de la « communauté internationale » à l'ONU), les États sont des unités spatiales très hétérogènes. Comme l'expliquent Pumain et Saint-Julien : « Entre la Chine et Pitcairn, les États du monde s'échelonnent de plus d'un milliard à quelques dizaines d'habitants, leurs superficies sont tout aussi incommensurables. Les écarts de dimension sont tels qu'ils rendent la comparaison difficile et peut-être illusoire, l'inégalité quantitative devenant telle qu'elle introduit une différence qualitative entre des unités spatiales qui n'ont plus en commun que quelques éléments de définition. » (Pumain & Saint-Julien, 1997, p. 19).

Pour plus de précisions et quand les besoins analytiques ou administratifs le justifient, d'autres niveaux sont donc parfois légitimement préférés ou mobilisés en complément au niveau national. Par exemple, pour étudier la géographie de la production scientifique des États-Unis ou du Canada, les États fédérés ou les provinces sont des niveaux qu'il est assez naturel de considérer, compte tenu de la compétence dont ils disposent dans

l'administration de l'enseignement supérieur et de la recherche. À l'échelle de l'Union Européenne, les régions ont également un rôle à jouer puisqu'elles bénéficient des soutiens financiers délivrés dans le cadre de la politique européenne de cohésion territoriale. Mais comme les régions administratives ne constituent pas une population d'entités spatiales homogènes et comparables à l'échelle européenne, le service statistique de l'Union Européenne, appelé Eurostat, a mis au point la « nomenclature des unités territoriales statistiques » ou NUTS. Il s'agit d'une grille d'équivalence entre niveaux administratifs infranationaux. Les rapports réalisés par Science-Metrix en 2013 pour la Commission Européenne témoignent de l'intérêt opérationnel de cette nomenclature pour notre sujet puisque des données sur la production scientifique y sont disponibles à la fois au niveau NUTS 2 (équivalent des régions françaises) et au niveau national (Campbell, Lefebvre, *et al.*, 2013). De même, beaucoup d'études en géographie de l'innovation se font au niveau de résolution NUTS 2. Ceci-dit, les spécialistes de l'innovation ne semblent pas toujours conscients des écueils liés à la nomenclature européenne sur lesquels Marie-Noëlle Comin revient dans sa thèse (Comin, 2009, p. 191), et notamment du fait que cette nomenclature met sur le même plan des divisions administratives fondées sur des critères distincts selon les pays. En plus, tandis que les frontières des États sont sensibles aux événements géopolitiques, les frontières régionales ne sont pas exemptes de recompositions au gré des réformes territoriales.

Plus préoccupant encore, il arrive que le niveau régional soit mobilisé par défaut, faute de pouvoir produire une analyse à un niveau spatial plus précis. En effet, dans un recueil réalisé par les spécialistes français de la proximité Alain Rallet et André Torre intitulé *Quelles proximités pour innover ?*, Jean-Alain Héraud *et al.* justifient leur méthode de la façon suivante : « L'idéal serait de travailler empiriquement sur des territoires construits au préalable, qui présentent les caractéristiques souhaitées de cohérence spatiale et organisationnelle mais les contraintes statistiques des données disponibles pour notre travail imposent le choix de la région au sens institutionnel comme variable proxy de territoire. » (Héraud, Munier, & Ronde, 2006). S'ils avouent renoncer à un idéal en faisant le choix de la région, ils compensent : « Comme par ailleurs, en matière de système d'innovation, la dimension de la gouvernance des territoires est loin d'être neutre et que les régions françaises sont amenées à jouer un rôle croissant dans les politiques d'innovation (voire de recherche), ce choix n'est pas exagérément arbitraire. » (Héraud *et al.*, 2006). Entre autres intérêts, cette citation confirme qu'il est préférable que l'échelon administratif mobilisé pour l'analyse dispose d'une compétence ou d'un caractère structurant dans le domaine socio-économique étudié. Par ailleurs, elle insiste sur la notion de « territoire » dans la mesure où l'entité spatiale d'analyse retenue serait d'autant plus pertinente qu'elle

fait sens dans la pratique quotidienne et les représentations des individus ou groupes qui en relèvent.

À cet égard, alors que la mesure de la production scientifique se fait encore essentiellement au niveau des États ou des régions, il apparaît souhaitable, même si techniquement plus délicat, de saisir la répartition des activités scientifiques à un niveau plus local. C'est en tout cas le point de vue de la géographie urbaine tel qu'il est exprimé par Myriam Baron, Denis Eckert et Laurent Jégou : « les approches "régionales", (...) constituent certes une bonne "approximation" quand elles coïncident uniquement avec les plus grandes villes scientifiques. Mais ce n'est pas le cas pour les espaces "régionaux" qui abritent plusieurs villes contribuant de manière équivalente à la production scientifique et dont le fonctionnement peut être plutôt polycentrique ; ni pour les espaces "régionaux" qui abritent des centres de production scientifique d'importance très différente. Dans cette dernière configuration, il convient de cerner non seulement le poids respectif des différentes classes de pôles scientifiques mais aussi d'identifier leurs relations. Tout ceci plaide pour des études, des analyses sur un système urbain dans son ensemble, voire sur les différents systèmes urbains. » (Eckert *et al.*, 2013). Comme le suggère la citation de Héraud *et al.* (*op. cit.*), ce point de vue est partagé par certains spécialistes de l'innovation, même s'ils ne disposent pas forcément des moyens pour le mettre en œuvre. Pour leur part, Héraud *et al.* préféreraient des analyses à un niveau plus fin que le niveau régional parce qu'ils considèrent qu'il existe un niveau plus local où se joue l'action, et au sein duquel s'échangent les idées et les compétences. Ces spécialistes de l'innovation sont sensibles au fait que les murs des laboratoires ne sont pas hermétiques, que les scientifiques ne sont pas coupés de leur environnement immédiat et participent à un système local. L'approche par les « systèmes locaux d'innovation », évoquée au cours de l'état de l'art, tend à confirmer cette hypothèse (voir Partie 1. Chapitre 3.1.1). Les paragraphes qui suivent visent à synthétiser les résultats connus pouvant expliquer la densité relationnelle se trouvant au fondement de la notion de « systèmes locaux d'innovation ».

Pour comprendre la cohésion des « systèmes locaux » du point de vue des sciences et de l'innovation, la nouvelle sociologie économique fournit des pistes intéressantes. Issue des études de Marc Granovetter sur les recouvrements entre le marché et les réseaux personnels, cette sociologie, dont se nourrit le Groupe « dynamiques de proximité » déjà évoqué, montre que la genèse des relations professionnelles est sujette à des effets d'encastrement. En particulier, en analysant la dynamique des relations industrielles au sein de laboratoires de recherche, Jacqueline Estades, Pierre-Benoît Joly et Vincent Mangematin ont trouvé que « l'espace où se produisent et s'utilisent les connaissances scientifiques et techniques n'est pas uniforme et amorphe. Il est structuré par des liens de proximité, des relations qui se tissent au gré de l'histoire et des décisions des acteurs. »

(Estades, Joly, & Mangematin, 1996). En s'appuyant sur le travail de ces derniers, Michel Grossetti et Marie-Pierre Bès ont distingué trois logiques de « mise en contact » permettant de construire une collaboration science-industrie (cette typologie est transposable à d'autres types de partenariats) : la logique de réseau, la logique d'institution et la logique de « marché » (Grossetti & Bès, 2002). La première relève d'une relation établie au préalable dans le cadre, bien souvent, des milieux professionnels ou de formation fréquentés dans leur parcours par les partenaires ; la seconde est déterminée par la présence d'une organisation intermédiaire qui met en contact des membres du laboratoire et de l'entreprise ; la dernière découle de la possibilité d'entrer en contact *via* les informations publiques disponibles sur chacun (par exemple, les coordonnées d'un auteur sur une publication scientifique) ou dans des événements tels que les salons, foires ou colloques. Plus généralement, le réseau social personnel ou réseau de l'ensemble des personnes de notre entourage à qui l'on peut faire appel pour se procurer une ressource – celui sur lequel s'appuie la logique de réseau – évolue dans le temps et l'espace fréquenté, au fur et à mesure des transitions (ou passage d'un âge de la vie à un autre) et des cercles sociaux fréquentés (Bidart, 2008) ; mais étant donné le périmètre géographique de notre quotidien et l'organisation de notre activité professionnelle, une partie importante des membres de ce réseau est susceptible d'appartenir à notre voisinage physique immédiat (de l'ordre d'une agglomération urbaine ou d'un bassin d'emploi) ; ce que confirment les analyses sociologiques de réseaux personnels (Grossetti, 2005).

Parallèlement, s'il est admis que les rencontres physiques sont utiles pour entretenir ou déclencher des relations, plusieurs études démontrent qu'elles ne sont pas indispensables pour rester en lien. Pour André Torre, la « proximité géographique temporaire » est suffisante pour maintenir une relation (Torre, 2008) ; de même que pour Michel Grossetti et Marie-Pierre Bès, « le travail en collaboration n'est pas particulièrement plus difficile à distance à condition que certains participants puissent se déplacer sans problème, et que les dispositifs techniques et organisationnels soient adaptés à cette situation. » (Grossetti & Bès, *op. cit.*). Dès lors, pour des raisons structurelles, la plupart des relations interpersonnelles sont susceptibles de se développer dans le milieu local (expliquant la cohésion observée à ce niveau) mais, dans certaines conditions, il est possible, voire nécessaire, d'entretenir ou d'activer des relations, notamment professionnelles, à distance. Ainsi, le réseau de collaborations ou l'« univers de références » scientifiques d'un chercheur à un instant donné est lié à sa trajectoire non seulement professionnelle (sa spécialité scientifique) mais aussi géographique (les différents lieux de recherche qu'il a fréquentés et avec lesquels il continue parfois d'interagir) (voir Partie 3. Chapitre 11.).

Beaucoup de dispositifs de coordination sont toutefois plus opérants au niveau local si bien que certaines relations ont tendance à conserver un ancrage plus local que d'autres.

En particulier, les relations « maître-élève » et les relations « science-industrie » tendent à être plus locales parce qu'elles sont « inter-institutionnelles ». Comme les institutions de recherche ne sont pas coupées des besoins techniques, des besoins de formation et des préoccupations de la société environnante, elles tissent des liens avec le territoire dans lequel elles s'inscrivent. C'est d'autant plus net quand il existe des dispositifs administratifs qui encouragent ou facilitent localement l'établissement de ces relations. C'est par exemple la fonction des pôles de compétitivité, dispositifs étudiés par Antoine Grandclément (2012) ou encore la fonction d'encadrement des thèses regardée de près par Olivier Godechot puis par Bastien Bernela *et al.* (Godechot et Louvet, 2010 ; Godechot, 2012 ; Bernela, Bouba-Olga et Ferru, 2013) ; dispositifs qui mettent plus spontanément en relation des individus appartenant aux mêmes systèmes urbains. Ces logiques systémiques expliquent l'idée, souvent mobilisée dans les discours territorialisés sur la recherche, que les sciences les plus appliquées ont un potentiel de développement local plus fort. Comme Roderik Ponds l'a montré à partir de données sur les collaborations scientifiques néerlandaises, la portée des échanges de savoirs varie en fonction de la nature des relations et du moment où elles sont mises en œuvre (Ponds, 2008). D'après lui, il faut que de nombreuses conditions soient réunies pour que les collaborations scientifiques se déploient à longue distance. Dans son analyse, même si la majeure partie des co-publications sont limitées à un périmètre local, les co-publications les plus insensibles aux frontières géographiques sont celles qui se déploient au sein du milieu académique. S'appuyant sur les théories de la proximité, il en déduit que la proximité cognitive et institutionnelle qui règne entre les acteurs du monde académique rend plus accessoire et contingent le rôle de la proximité géographique. Sa thèse suggère que les collaborations entre établissements de recherche et entreprises (inter-institutionnelles) se déploient moins facilement à longue distance que les collaborations internes au monde académique (intra-institutionnelles). Ces dernières sont sans doute d'autant plus transfrontalières qu'elles sont encouragées par des dispositifs comme les PCRD (Programme Cadre de Recherche et Développement européens). Enfin, Ponds suggère que la fréquence et la portée de ce dernier type de collaborations peut dépendre du stade où se trouve la recherche, ce que nous avons vérifié à travers les résultats présentés dans la Partie 3. Chapitre 11.

Dès lors, tandis que la remarquable densité des échanges scientifiques locaux est démontrée par les sciences régionales et urbaines, on connaît moins bien la distribution, la portée et la dynamique des flux scientifiques entre systèmes locaux ou urbains — même s'il existe quelques travaux pionniers comme ceux des géographes scandinaves (Å. E. Andersson & Persson, 1993). Cet argument explique que nous ayons choisi de considérer les systèmes locaux ou agglomérations urbaines comme niveau élémentaire d'analyse pour étudier la géographie de la production et des collaborations scientifiques à l'échelle mon-

diale, sans nous préoccuper outre mesure de la synergie interne à ces entités — qui est de toute façon délicate à estimer à cette échelle étant donné les limites déjà évoquées concernant l'hétérogénéité des organisations scientifiques.

À partir du moment où l'on souhaite prendre le niveau local comme niveau élémentaire d'analyse, il reste à trouver une solution permettant d'associer l'information bibliographique à des entités spatiales locales, autrement dit de ventiler la production scientifique par système local ou agglomération. En effet, beaucoup de spécialistes de l'innovation se contentent d'analyses au niveau régional parce qu'il n'existe pas de population d'entités spatiales correspondant au niveau du système local ou urbain qui soit définie de façon homogène à l'échelle mondiale. C'est pour cette raison qu'il a fallu mettre au point une méthode permettant de délimiter des « agglomérations scientifiques » en tenant compte de la répartition mondiale de la population, de la production scientifique et des activités humaines. Cela n'a été possible que parce qu'on a procédé, en premier lieu, au « géocodage intégral des lieux de production de la science » (Eckert *et al.*, *op. cit.*). En codant les données bibliographiques au meilleur degré de précision géographique atteignable, on s'est donné l'opportunité de procéder à des agrégations spatiales successives :

- d'abord au niveau des villes ou des agglomérations, puis à des niveaux supérieurs (région urbaine ou administrative, pays, zone transfrontalière, etc.) ;
- et d'arbitrer en fonction des besoins de la recherche sans être prisonnier des découpages administratifs.

Cette démarche nous a permis de concilier « une analyse globale du fonctionnement géographique de la science et des études multiniveaux. » (*ibid.*).

1.3. Le géocodage des données bibliographiques

Avant de pouvoir agréger les données de publication au niveau d'agglomérations ou de systèmes locaux définis à l'aide d'une procédure unifiée permettant la comparaison à l'échelle mondiale, chaque adresse de publication a été associée à un point correspondant aussi précisément que possible à la localité déclarée par les auteurs dans leur adresse institutionnelle. La répartition de ces points a ensuite servi de repère pour définir le contour des agglomérations qui ont servi d'entités élémentaires d'analyse.

Lorsqu'ils signent, les auteurs de publications sont supposés indiquer leur ville d'appartenance dans leur adresse institutionnelle. De façon systématique, l'ISI Thomson Reuters soumet ces adresses à des traitements automatiques permettant de réordonner et d'autonomiser plusieurs champs : organisation principale, ville, province, pays. C'est que les auteurs de publications scientifiques renseignent leur affiliation comme bon leur semble, si bien que l'adresse qu'ils déclarent est intégrée au *Web of Science* sous la forme

d'un champ libre ou chaîne de caractères qui doit impérativement être structuré pour pouvoir être ré-exploité. Pour établir des statistiques, les algorithmes de l'ISI Thomson Reuters réordonnent et identifient des sous-chaînes à l'intérieur des adresses, tâchant de les associer au niveau de la hiérarchie institutionnelle ou géographique auxquelles elles correspondent et de les homogénéiser. Cependant, le travail d'homogénéisation réalisé par l'ISI est surtout axé sur la normalisation des champs « pays » et « organisation principale », qui sont les champs à partir desquels l'organisme commercialise des indicateurs et propose des analyses, ce qui explique que les champs « ville » et « province » stockés par l'ISI sont beaucoup moins exploitables en l'état. Ces derniers sont pourtant fondamentaux pour localiser précisément les données de production scientifique.

Nota bene : le contenu des triplets « ville-province-pays » est disponible dans les enregistrements de la base de données obtenus par voie de contractualisation auprès de l'ISI, mais le simple utilisateur de la base de données en ligne n'y a pas accès lorsqu'il réalise des extractions ponctuelles depuis l'interface *Web*. Ainsi, lorsqu'on extrait un corpus depuis la base de données en ligne – comme nous avons été amenée à le faire pour plusieurs ensembles de publications scientifiques portant sur la réparation de l'ADN, voir Partie 3 – il faut soi-même concevoir des routines automatiques pour décomposer les chaînes de caractères du champ « adresse » et en extraire les triplets « ville-province-pays ». Une autre solution – expérimentée par la même occasion – consiste à injecter directement le contenu intégral du champ adresse dans un outil de géocodage automatique pour lui laisser le soin de localiser et d'affecter un jeu de coordonnées géographiques (latitude, longitude) à chaque adresse (Leydesdorff & Persson, 2010). Favorables à cette solution, Leydesdorff et Persson ont mis au point des utilitaires permettant de visualiser immédiatement les résultats obtenus par ce biais sur *Google Earth* (format « .kml »)⁴⁸. Cependant, cette dernière approche génère des complications puisque le degré de précision que le « géocodeur » est en mesure d'atteindre varie d'une adresse à l'autre.

Par exemple, pour l'adresse « Université de la Sorbonne, Paris », un « géocodeur » propose plusieurs couples de coordonnées à l'intérieur de Paris mais aussi un couple de coordonnées à Nogent-sur-Marne ; parmi lesquels il n'est pas capable de trancher parce que l'université est multi-localisée⁴⁹. D'autres fois, lorsque le « géocodeur » ne connaît pas la localisation de l'organisation scientifique (par exemple « Université Toulouse 2, Albi »), il renvoie les coordonnées géographiques de la localité, ou, à défaut, de la province ou du pays. Or, dans la perspective d'harmoniser l'information spatiale dont on dispose pour chaque adresse institutionnelle, mieux vaut privilégier un seul niveau de

⁴⁸ Source : URL : <http://www.leydesdorff.net/software.htm>

⁴⁹ Résultat d'un test réalisé le 15/03/2015 en utilisant les services du site internet : <http://www.gpsvisualizer.com/geocode>

résolution auquel associer un seul point de publication. Cela limite aussi les risques d'erreurs puisque le nom de l'établissement peut se confondre avec un nom de rue, de localité ou de région, générant des résultats erronés. Ajoutons à cela qu'en ne faisant pas l'effort de structurer et de sélectionner l'information avant de la géocoder, la quantité d'information à traiter est bien plus conséquente (les adresses entières peuvent être très longues) et la procédure automatique prend alors beaucoup plus de temps, ce qui pose problème à partir du moment où, comme cela a été le cas dans le cadre de la « tâche 2 » du programme « Géoscience », on entend géocoder l'ensemble des adresses de publications indexées dans la base de données à plusieurs dates.

De toute façon, au début du travail collectif de géocodage des données du WoS (« tâche 2 »), on ne disposait pas du champ adresse dans son intégralité mais seulement des triplets « ville-province-pays » qui en ont été extraits (voir Partie 1. Chapitre 5.1.2). Les données géographiques étaient même « pré-nettoyées » puisque les scientomètres de l'OST de Montréal, avec lesquels nous avons collaboré, avaient déjà opéré des améliorations pour réaliser leurs propres analyses. En particulier, le champ province pour le Canada avait été désambiguïsé et homogénéisé. À partir de là, il s'agissait de localiser le plus finement possible à l'échelle mondiale les localités indiquées par les triplets associés aux publications du SCI Expanded en 1978, 1988, 1998 et 2008. Dans cette perspective, on s'est limité aux triplets associés à des quantités suffisantes de publications : n'ont été traités que les triplets associés à au moins deux publications à au moins une des dates étudiées. Mais, y compris pour coder les combinaisons les plus fréquentes, il a fallu accomplir un travail important de désambiguïstation du champ « ville ». En effet, nombreux sont les triplets qui désignaient la même localité avec des graphies différentes ou bien des erreurs d'orthographe. En plus de la liberté d'écriture des scientifiques, des erreurs récurrentes avaient pu se glisser dans la base lors de la saisie ou la manipulation numérique des données.

De ce fait, un premier travail de normalisation de l'information a été nécessaire, qui consistait à s'appuyer sur le contenu des catalogues de toponymes ou *gazetteers* comme Geonames⁵⁰ ou Nominatim⁵¹ et à supprimer les erreurs manifestes de saisie (Jégou, 2014). À partir de là, un grand nombre de triplets, surtout les plus communs, ont pu être corrigés puis géolocalisés sans trop de mal. Pour localiser les triplets résistants, il ne restait pas d'autres choix que de les soumettre, en l'état, aux solutions de géolocalisation automatique. Le moyen technique actuel le plus commode pour géocoder un nombre important de lieux consiste à mobiliser les services de géocodage automatique en ligne, proposés par des sociétés comme Google ou Yahoo ! Ces services interrogent leurs sources de données

⁵⁰ Source : URL : <http://www.geonames.org/>

⁵¹ Source : URL : <http://nominatim.openstreetmap.org/>

cartographiques pour associer un jeu de coordonnées géographiques unique (une latitude et une longitude) à chaque toponyme qu'on leur soumet.

De tels services de géocodage ont été mobilisés chaque fois que notre équipe s'est procurée de nouveaux extraits de la base de données. Les principales vagues de géocodage sont résumées dans le Tableau 9 qui en distingue trois : Printemps 2010 (données OST Montréal, années : 1978, 1988, 1998, 2008) ; Printemps 2012 (données OST Paris, années : 1999- 2001 ; 2006-2008) ; Printemps 2015 (données OST Paris, années 2002-2005 ; 2009-2012). Pour cela, les services de l'API⁵² *Google Maps*, globalement les plus satisfaisants, ont été privilégiés en les complétant parfois par d'autres solutions. Puisqu'il y a eu plusieurs étapes de vérification du géocodage et d'harmonisation des triplets au fur et à mesure du travail, le protocole est synthétisable sous la forme d'une chaîne (Figure 19), reprise d'une présentation de Laurent Jégou à la Conférence de Géographie de l'innovation d'Utrecht (Jégou, 2014). Comme on le voit sur cette figure, les allers-retours ont été nombreux entre le codage géographique des données et la correction des données.

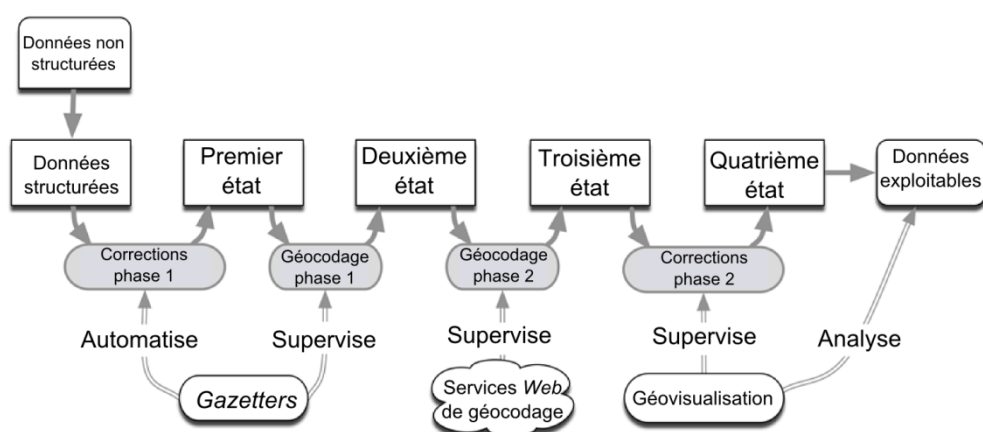


Figure 19 – La chaîne de structuration des données géographiques. Schéma repris et traduit du proceeding « *Towards spatially referenced academic data at global scale : the full geocoding of Wos-Datasets, methods and results* » (Jégou, 2014)

En appui à ce travail, Laurent Jégou a conçu un dispositif permettant, entre autre, d'explorer les résultats du géocodage. Il s'agit d'un logiciel de géovisualisation construit avec la bibliothèque de fonctions OpenLayers à partir duquel il est possible de vérifier si le jeu de coordonnées géographiques attribué à chaque triplet est correct et, sinon, de proposer une solution de localisation plus juste ou plus précise.

⁵² API : *Application programming interface*, ensemble organisé de fonctions disponibles pour la programmation.

Parmi ses nombreuses fonctions, cet utilitaire rend compte de la valeur de l'indice de précision associé aux résultats du géocodage. En effet, lors de la première et seconde vague de géocodage, chaque réponse donnée par l'API *Google Maps* s'accompagnait d'une valeur comprise entre 0 et 10 : le niveau d'*accuracy* (il valait 2 pour le niveau régional et au moins 4 pour le niveau local). Grâce à cette information (qui n'est plus donnée par la version actuelle de l'API), on pouvait facilement repérer des « faux positifs », c'est-à-dire des points de publication que le « géocodeur » prétendait avoir bien placé (bon niveau d'*accuracy*) sans qu'ils correspondent au résultat recherché. C'est par exemple le cas d'un point « Guadeloupe » localisé à la périphérie de Saint-Romans et Saint-Marcellin en Isère. Pour ce point « Guadeloupe », le « géocodeur » indique avoir trouvé un résultat d'une très grande précision (*accuracy* = 6). Cette valeur est en réalité alarmante car lorsque le jeu de coordonnées trouvé correspond à celui d'une localité, le niveau d'*accuracy* est de 4. Cela signifie que c'est une rue, un hameau ou un établissement répondant au nom de « La Guadeloupe » qui a été localisé à cet endroit et non une localité. Mais, comme en attestent les données de production scientifique associées aux points de publications (Figure 20), cette erreur est négligeable puisqu'elle ne concerne qu'une dizaine de publications (2 en 1998 et 9 en 2008)

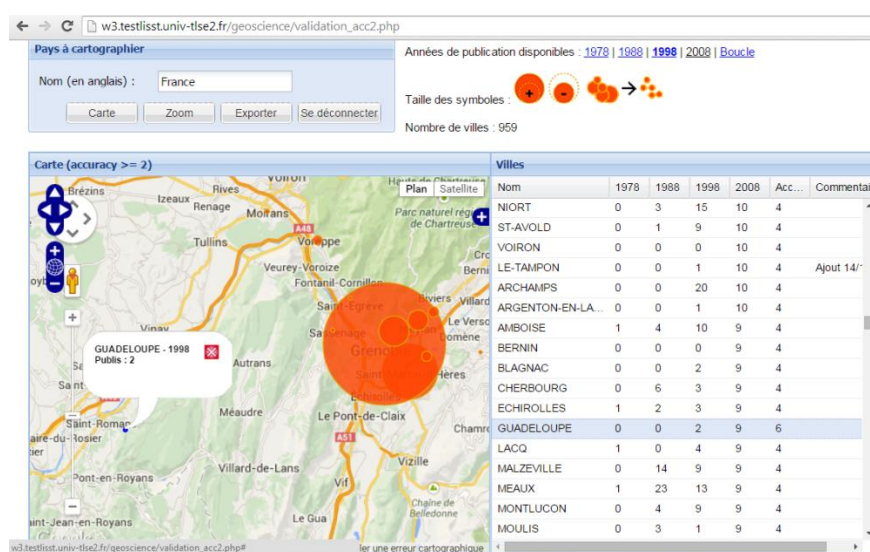


Figure 20 – L'outil d'aide à la vérification du géocodage. Capture d'écran témoignant d'une erreur de géolocalisation : un point « Guadeloupe » dans l'Isère, France

Bien que chaque triplet devrait pouvoir être associé à une seule localité, il n'est pas toujours évident de déterminer laquelle est désignée, en raison des possibles homonymies (plusieurs villes portent le même nom), mais aussi des erreurs ou des données manquantes (le champ province est rarement renseigné pour les données antérieures à 2008, et il peut

y avoir des inversions entre les différents champs). Ainsi, lors d'une vérification systématique réalisée à l'aide du logiciel de géovisualisation pour le cas de la France, j'ai recensé plusieurs types d'approximations et d'imprécisions avec lesquelles il faut encore composer une fois que la procédure de géocodage et de vérification est terminée.

Voici les principales d'entre elles :

1) Il y a de nombreux homonymes parmi les triplets, autrement dit, plusieurs solutions (jeux de coordonnées géographiques) possibles pour un même triplet : par exemple, Saint-Denis de l'Ile de la Réunion et Saint-Denis en Ile-de-France sont parfois désignés par des triplets identiques. Puisque pour le cas de la France, il est rare que le champ « province » soit renseigné, l'information « Saint-Denis – France » n'est pas d'une précision suffisante pour être affectée avec certitude à l'une ou l'autre des localités. Autrement dit, il serait faux d'affecter toutes les contributions scientifiques issues de « Saint-Denis – France » à l'un des Saint-Denis plutôt qu'à l'autre. Pour un tel cas de figure, il n'y a qu'en disposant des adresses entières ou encore du code postal que nous pouvons espérer trancher, et nous n'en disposons pas pendant la première phase du projet.

2) Il n'est parfois pas possible d'associer un triplet à une localité précise puisque il existe des lieux dits, comme le quartier de la Défense (région parisienne), le parc technologique de Sophia-Antipolis (région niçoise) ou celui de Courtabœuf (Essonne), qui s'étendent sur plusieurs municipalités.

3) Enfin, il arrive que la ville ne soit pas renseignée (il existe des cas, notamment pour les publications du corps militaire, où aucun lieu géographique n'est déclaré) ou très mal (une faute de frappe ou une erreur de translittération) si bien qu'à moins d'avoir trouvé une correction, le géocodage se fait au niveau régional (c'est d'autant plus probable si la région ou province a été renseignée par erreur dans le champ « ville ») ou encore national (ce qui rend le résultat inexploitable).

Une fois en possession des données issues de l'OST Paris (2^{ème} et 3^{ème} vague de géocodage), nous avons pu approfondir les méthodes de vérification puisque nous disposons, en plus des triplets « ville-province-pays », du champ « adresse » au complet et d'un champ « code postal ». Ces informations supplémentaires se sont parfois révélées utiles pour trancher en cas d'homonymie. Dans cette perspective, j'ai travaillé avec Laurent Jégou à perfectionner le codage des adresses de la France et de l'Allemagne, et Fabien Goblet a travaillé à l'automatisation des traitements pour les pays restants. Grâce à l'application de scripts visant à homogénéiser l'information spatiale, la qualité des résultats du géocodage pour tous les pays a connu de notables améliorations.

En fait, à chaque livraison de données, le géocodage et le « nettoyage » des nouveaux triplets ont permis d'enrichir le « fichier recette » c'est-à-dire le fichier en constante amélioration qui, pour chaque « triplet-source », indique un « triplet-destination »

(éventuellement correctif du précédent) et un jeu de coordonnées géographiques. Comme le résume le tableau suivant (Tableau 9), les « triplets-destination » sont moins nombreux parce qu'ils proposent une seule graphie par localité.

Vagues de géocodage	Années de publication	Fournisseur des données	Périmètre	Triplets origine (nb)	Triplets destinations (nb)	Nom de la recette
Printemps 2010 - Automne 2011	1978, 1988, 1998, 2008	OST Montréal	SCI Expanded	28 271	17 615	Recette 1 (Octobre 2011)
Printemps 2012 - Automne 2013	1999- 2001 ; 2006-2008	OST Paris	WoS	40 675	21790	Recette 13 (Octobre 2013)
Printemps 2015 - Été 2015	2002-2005 ; 2009-2012	OST Paris	WoS	88 507	33 204	Recette 15 (Mai 2015)

Tableau 9 – Bilan des trois vagues de géocodage : l'évolution du fichier recette

Ce « fichier recette » offre un moyen d'homogénéiser l'information géographique pour faciliter son traitement. Comme on peut le deviner à partir de la capture d'écran de la recette n°13 (Figure 21), si l'on souhaite appliquer avec succès ce fichier à une extraction brute de la base de données (réaliser une jointure), il est nécessaire de procéder à quelques prétraitements parmi lesquels une harmonisation de la casse et de la typographie des « triplets sources » (passage en majuscules non accentuées, remplacement des espaces par des tirets, etc.).

country_src	province_src	city_src	country_dest	province_dest	city_dest	lat	lng
ENGLAND	BED	SHARNBROOK	UNITED-KINGDOM	ENGLAND	SHARNBROOK	52.2261	-0.544336
ENGLAND	BED	SILSOE	UNITED-KINGDOM	ENGLAND	SILSOE	52.0085	-0.424755
ENGLAND	BEDFORD	SHARNBROOK	UNITED-KINGDOM	ENGLAND	SHARNBROOK	52.2261	-0.544336
ENGLAND	BEDFORDSHIRE	BEDFORD	UNITED-KINGDOM	ENGLAND	BEDFORD	52.1357	-0.468044
ENGLAND	BEDFORDSHIRE	CRANFIELD	UNITED-KINGDOM	ENGLAND	CRANFIELD	52.0685	-0.608761
ENGLAND	BEDFORDSHIRE	DUNSTABLE	UNITED-KINGDOM	ENGLAND	DUNSTABLE	51.8876	-0.523129
ENGLAND	BEDFORDSHIRE	LUTON	UNITED-KINGDOM	ENGLAND	LUTON	51.8797	-0.417558
ENGLAND	BEDFORDSHIRE	SANDY	UNITED-KINGDOM	ENGLAND	SANDY	52.1298	-0.289933
ENGLAND	BEDFORDSHIRE	SHARNBROOK	UNITED-KINGDOM	ENGLAND	SHARNBROOK	52.2261	-0.544336
ENGLAND	BEDFORDSHIRE	SILSOE	UNITED-KINGDOM	ENGLAND	SILSOE	52.0085	-0.424755
ENGLAND	BEDS	CRANFIELD	UNITED-KINGDOM	ENGLAND	CRANFIELD	52.0685	-0.608761
ENGLAND	BEDS	LUTON	UNITED-KINGDOM	ENGLAND	LUTON	51.8797	-0.417558
ENGLAND	BEDS	SANDY	UNITED-KINGDOM	ENGLAND	SANDY	52.1298	-0.289933
ENGLAND	BEDS	SHARNBROOK	UNITED-KINGDOM	ENGLAND	SHARNBROOK	52.2261	-0.544336
ENGLAND	BEDS	SILSOE	UNITED-KINGDOM	ENGLAND	SILSOE	52.0085	-0.424755
ENGLAND	BEDSFORDSHIRE	AMPTHILL	UNITED-KINGDOM	ENGLAND	AMPTHILL	52.0274	-0.49524
ENGLAND	BEDSHIRE	LUTON	UNITED-KINGDOM	ENGLAND	LUTON	51.8797	-0.417558
ENGLAND	BEFORDSHIRE	SHARNBROOK	UNITED-KINGDOM	ENGLAND	SHARNBROOK	52.2261	-0.544336
ENGLAND	BER	ASCOT	UNITED-KINGDOM	ENGLAND	ASCOT	51.4108	-0.674823
ENGLAND	BER	HARWELL	UNITED-KINGDOM	ENGLAND	HARWELL	51.597	-1.29299
ENGLAND	BER	MAIDENHEAD	UNITED-KINGDOM	ENGLAND	MAIDENHEAD	51.5228	-0.720209
ENGLAND	BER	READING	UNITED-KINGDOM	ENGLAND	READING	51.455	-0.969088
ENGLAND	BER	WOKINGHAM	UNITED-KINGDOM	ENGLAND	WOKINGHAM	51.411	-0.834926
BELGIUM	ZWI	GHENT	BELGIUM		GHENT	51.0536	3.72087
CZECH-REPUBLIC	JIN	PRAGUE	CZECH-REPUBLIC		PRAGUE	50.0878	14.4205
CZECH-REPUBLIC	KRC	PRAGUE	CZECH-REPUBLIC		PRAGUE	50.0878	14.4205
CZECH-REPUBLIC	MOTOL	PRAGUE	CZECH-REPUBLIC		PRAGUE	50.0878	14.4205
CZECH-REPUBLIC	SUC	PRAGUE	CZECH-REPUBLIC		PRAGUE	50.0878	14.4205

Figure 21 – Extrait du fichier recette n°13 permettant d'associer à un triplet source (ville-province-pays) un triplet destination et un couple de coordonnées géographiques

En s'aidant de ce fichier, réaliser des analyses spatiales au niveau de la localité (le point de géocodage) est possible et peut donner l'illusion d'une extrême précision mais ce niveau de finesse est très variable d'une zone géographique à une autre. Il dépend no-

tamment de la complexité des découpages administratifs qui diffèrent d'un pays à l'autre. De ce fait, se fier à l'information géographique ponctuelle pourrait contribuer à appauvrir la validité de l'analyse spatiale puisque les unités d'analyse seraient trop hétérogènes, incommensurables.

Cet argument statistique justifie pleinement le regroupement de plusieurs « points de publications » au sein d'agglomérations, mais il existe également, on l'a vu dans la section précédente, des arguments théoriques qui permettent d'étayer ce choix méthodologique, comme l'idée que le niveau local correspondant à la réalité de l'entreprise humaine quotidienne peut s'étendre à plusieurs localités, une ville centre et sa proche périphérie par exemple.

2. La délimitation des agglomérations scientifiques

« Le pivot des organismes urbains, c'est la ville. Nous avons vu quelles difficultés il y avait à la définir. Mais l'habitude a créé un certain vocabulaire qu'il semble bon d'explicitier. L'agglomération est la forme la plus simple du développement urbain. Elle se définit, classiquement, comme une ville entourée d'une banlieue. » (Beaujeu-Garnier, 1980, p. 107-108)

Parmi les travaux actuels en scientométrie spatiale, la question du niveau d'agrégation des données est peu débattue. Profitant des nouvelles possibilités offertes par les outils de géocodage automatique des adresses, certains chercheurs considèrent que le meilleur niveau de précision obtenu par le géocodage est un niveau satisfaisant auquel agréger les données. En dehors du fait que, les adresses n'étant pas normalisées, leur niveau de précision est inégal, il reste un problème d'hétérogénéité spatiale. Puisque d'un pays à un autre le niveau de fragmentation administrative varie considérablement, le périmètre des entités administratives élémentaires ou « localités » est lui-même très variable. En particulier, au niveau de l'aire urbaine de Pékin (aussi appelée Beijing) et celle de Paris, le nombre de points de publications, c'est-à-dire le nombre de localités géocodées à partir des affiliations des chercheurs (normalement des municipalités), varie du simple au centuple. Cela vient du fait que la surface des deux municipalités est incomparable : 100 km² pour Paris et cent fois plus pour Pékin. La planche d'illustrations ci-dessous (Figure 22) extraite de l'outil de géovisualisation développé pour la vérification du géocodage et l'exploration des données montre la diversité des cas de figure en matière de répartition des points de publication.

Publications dans le WOS, localisations et zones urbaines, variété des référencements

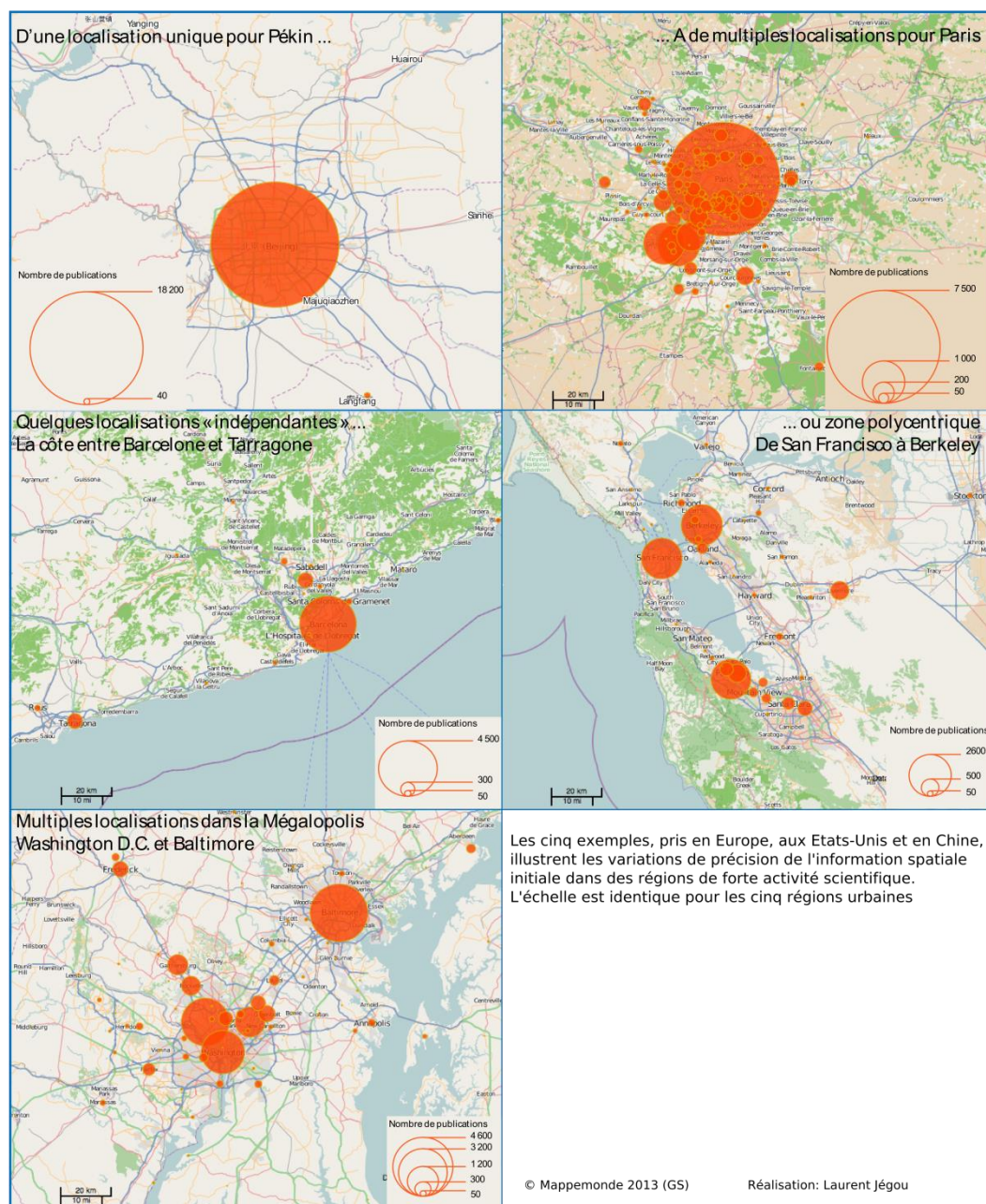


Figure 22 – Publications dans le Web of Science, localisations et zones urbaines, variété des référencements. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)

Puisque le niveau urbain est de plus en plus souvent pris comme niveau d'analyse scientométrique, il est essentiel d'attirer l'attention des spécialistes comme Loet Leydesdorff et Olle Persson sur cette question qu'ils ont tendance à négliger pour réaliser leurs mesures (Leydesdorff & Persson, 2010). La lettre aux éditeurs du *Journal of Infor-*

metrics rédigée par Lutz Bornmann et Andrew Plume en 2011 suggère que la préoccupation existe, mais ils ne font pas le tour de la question dans la mesure où ils interrogent l'effet de l'agrégation des données sur l'impact scientifique d'une seule agglomération : celle de Munich. Après avoir comparé les indicateurs d'impact scientifique obtenus pour la ville de Munich en harmonisant ou pas les adresses et en y agrégeant ou pas les publications écrites depuis les localités limitrophes (l'aire urbaine de Munich), les auteurs concluent que le fait d'inclure la production de Martinsried (en banlieue de Munich) pour mesurer l'activité de Munich n'a pas d'influence significative sur son impact : « Ces résultats indiquent que pour Munich la valeur d'impact mesurée à partir des citations ne dépend pas de la prise en compte des banlieues ou des petites villes à proximité ni des différentes variantes du nom de la ville en dehors de la dénomination principale München. »⁵³ (Bornmann & Plume, 2011). Certes, si le niveau de fragmentation administrative était le même partout, il y aurait moins d'inconvénients à prendre comme unité élémentaire d'analyse la plus petite division administrative et il ne serait même pas nécessaire d'avoir recours à une procédure désambiguïsation des toponymes, mais le niveau de fragmentation administrative est très loin d'être stable de sorte qu'il existe des endroits du monde où l'organisation spatiale n'a rien à voir, comme entre Moscou et la côte Est américaine : « Dans un espace de quelques dizaines de kilomètres carrés, bien plus petit que la municipalité de Moscou, on trouve trois centres majeurs : la ville de Washington DC proprement dite (2 500 publications), mais avant elle le grand centre universitaire de Bethesda dans le Maryland (3 178) et, un peu derrière, College Park (1 236) » (Grossetti *et al.*, 2013).

Les géographes sont habitués à ce problème d'hétérogénéité lié à l'inégale fragmentation administrative de l'espace. En effet, il existe en géographie de nombreux travaux qui rendent compte des différentes définitions de la ville, d'une part, et des différentes façons de délimiter des aires ou systèmes urbains, d'autre part. Ces travaux envisagent généralement la nécessité de se donner une définition homogène de l'objet « ville » pour réaliser des comparaisons internationales (Moriconi-Ebrard, 1991).

2.1. Définir la ville et délimiter des zones urbaines

Prétendre réaliser une analyse spatiale au niveau des villes est une revendication qui n'est claire qu'en apparence parce que le mot « ville » est utilisé pour désigner des portions d'espace de taille (en superficie et en nombre d'habitants) et de nature très di-

⁵³ L'original : « These results indicate that for Munich the citation impact values do not depend on the consideration of suburbs or small cities, respectively, in close proximity and different name variants besides the variant München. » (Bornmann & Plume, 2011).

verses. Au début des années 1990, dans le cadre de travaux doctoraux dirigés par Denise Pumain, Céline Rozenblat et Nadine Cattan (pour l'Europe) et François Moriconi-Ebrard (pour le monde) ont répertorié la variété des définitions des villes dans la statistique publique d'un pays à l'autre (Cattan & Rozenblat, 1991 ; Moriconi-Ebrard, 1991).

Dans beaucoup de pays, les services statistiques déterminent quelles unités spatiales peuvent être qualifiées d'« urbaines ». Cattan et Rozenblat expliquent que « pour les pays où cette définition existe, elle se réfère toujours à une unité administrative de base, en général la municipalité, parfois à une subdivision de cette unité de base. » (*ibid.*). Par ailleurs, Moriconi-Ebrard remarque que : « peu de concepts sont aussi disparates dans le monde que celui de "population urbaine". Cette diversité traduit différents points de vue sur la ville. » (*ibid.*). À la suite d'un vaste travail de recherche documentaire à travers les services statistiques d'un maximum de pays, ce dernier distingue 5 approches du fait urbain :

1. Une localité est considérée comme urbaine si elle est définie comme telle par des fonctions administratives (critère administratif)
2. Une localité est considérée comme urbaine si le nombre d'habitants de cette localité dépasse un certain seuil (critère quantitatif)
3. Une localité est considérée comme urbaine si elle n'est pas rurale, c'est-à-dire si le secteur agricole y est peu représenté (critère socio-économique)
4. Une localité est considérée comme urbaine d'après la quantité et la qualité des emplois, des services, des fonctions et des équipements qu'elle possède (critère fonctionnel)
5. Une localité est considérée comme urbaine compte tenu d'indicateurs démographiques comme la densité de population, les classes d'âge ou le taux de masculinité. (critère démographique)

Dans les faits, les meilleures définitions sont les définitions mixtes qui tiennent compte de plusieurs critères à la fois mais, d'après Moriconi-Ebrard, ce sont le luxe des petits pays et des pays riches par opposition aux définitions en vigueur dans les pays les moins équipés en main d'œuvre statistique. C'est ainsi que le critère le plus répandu dans le monde est également le plus simple à mobiliser : il s'agit du critère quantitatif reposant sur le nombre d'habitants.

C'est ce critère basique que Moriconi-Ebrard a choisi de retenir pour sélectionner l'ensemble des villes qui peuplent la base des grandes villes du monde GÉOPOLIS, dont il est l'artisan. Dans cette base de données, sont appelées « villes du monde », les villes supérieures à 10 000 habitants. Mais, pour que les comparaisons internationales soient pertinentes, il applique ce critère à des « agglomérations morphologiques » et non aux « unités locales » dont la taille et la nature diffèrent excessivement d'un pays à l'autre. Pour ce

faire, encore faut-il pouvoir estimer la population urbaine au niveau d'entités cohérentes d'un point de vue morphologique, autrement dit caractérisées par un bâti continu.

Ce travail est d'autant plus délicat que : « les limites administratives des villes coïncident rarement avec celles de l'agglomération morphologique. Le front des constructions se déplace sans cesse, si bien que les limites administratives ne peuvent les suivre. Tout le problème est donc de traduire ces limites morphologiques en termes de limites administratives, les seules sur lesquelles on dispose de données statistiques » (*ibid.*). Le défi de Moriconi-Ebrard revient à compiler des données de population produites à des niveaux inférieurs, ou dans quelques cas exceptionnels à des niveaux supérieurs, pour obtenir des indicateurs statistiques valables au niveau des « agglomérations morphologiques » délimitées à partir d'informations sur la continuité du bâti. Ce travail a été pionnier puisque « GÉOPOLIS est le premier véritable inventaire des grandes agglomérations. Il n'existait jusqu'alors aucune source capable de donner des informations de base telles que la localisation, le nombre d'habitants, le taux de croissance des plus grandes villes du monde. » (*ibid.*).

En revanche, le problème de la délimitation de zones urbaines n'est pas nouveau. Son intérêt a été identifié par les services statistiques de plusieurs pays dès les années 1950. En 1960, Jean Bastié et Marcel Brichler écrivent dans la revue *Population* de l'Institut National d'Études Démographiques (INED) au sujet de la « nécessité de délimiter des agglomérations urbaines » (Bastié & Brichler, 1960). L'étude menée par ces auteurs sur la zone urbaine de Paris est présentée comme « une contribution à l'élaboration d'une méthode internationale dont le besoin se fait sentir. » (*ibid.*). La continuité du bâti n'est qu'un des nombreux critères envisagés à cet égard. Il est utilisé pour délimiter les deux premiers périmètres urbains autour de Paris, à savoir l'« agglomération restreinte » et la « couronne urbaine ». En revanche, la « couronne suburbaine » et la « zone d'attraction de l'agglomération » ne remplissent pas ce critère ou faiblement. Les communes englobées dans ces deux derniers périmètres sont, pour leur part, caractérisées par la fréquence des échanges quotidiens avec la capitale dont témoignent les mouvements pendulaires et, en cela, remplissent des critères fonctionnels (Figure 23).

D'après ces auteurs, dans le cas des très grandes villes, la variété des exploitations statistiques exige de définir plusieurs périmètres : « il ne semble pas possible de satisfaire des besoins très différents avec une seule limite ; on est amené à en définir plusieurs de plus en plus étendues ; ainsi, chacun peut choisir celle qui paraît s'adapter le mieux à son propre travail. » (*ibid.*). Cette remarque est toujours valable aujourd'hui où il est devenu commun de distinguer entre deux types complémentaires de zones urbaines : l'agglomération morphologique et l'aire urbaine fonctionnelle. En particulier, il est fréquent de déduire une aire fonctionnelle à partir d'une agglomération morphologique,

c'est-à-dire de prendre cette dernière pour noyau autour duquel gravitent les satellites de l'aire fonctionnelle. Outre le recours à des critères objectifs, il arrive aussi que les choix en matière de délimitation s'appuient sur des connaissances subjectives du terrain : on parle de « dire d'expert ».

1. Agglomération restreinte. (Ville de Paris plus couronne urbaine).	1 a. — Zone d'habitation collective. 1 b. — Zone d'habitation mixte (immeubles collectifs et maisons individuelles).	Très forte densité de population et très forte proportion d'immeubles collectifs. Continuité presque absolue des constructions.
2. Couronne suburbaine.	Zone d'agglomération plus lâche.	Forte densité associée à une continuité étroite des constructions.
3. Zone d'attraction de l'agglomération.	Zone dortoir non agglomérée.	Très faible proportion d'agriculteurs, forte expansion, population des communes et densité relativement élevées, mouvements migratoires quotidiens importants avec Paris, présence d'un moyen de transport fréquent avec Paris, faible continuité des constructions.
		Caractéristiques voisines de celles de la zone précédente, le caractère urbain étant cependant un peu moins accentué et sans la continuité des constructions.

Figure 23 — Les trois limites successives de l'agglomération parisienne en 1960. Schéma extrait de l'article « Délimitation de l'agglomération parisienne » (Bastie & Brichler, 1960)

Étant donné la multiplicité des découpages possibles, les nombreux services nationaux de statistiques qui se sont penchés sur ce problème dans les années 1970-1990 ont abouti à autant de définitions différentes des aires urbaines et donc à des entités trop hétérogènes pour la comparaison internationale : « La difficulté fondamentale d'une comparaison internationale tient au fait que les concepts, mis au point pour cerner cette nouvelle dimension territoriale de la ville par les services nationaux de statistiques, reposent sur des critères différents. » (Cattan & Rozenblat, *op. cit.*). Comme l'indique Joël Charre dans son ouvrage sur l'utilisation des statistiques en géographie : pour la validité des analyses spatiales, il est nécessaire de travailler avec une population d'entités spatiales homogènes, c'est-à-dire de même nature à l'échelle considérée (Charre, 1995). C'est ce qui justifie que dès les années 1970, l'ONU, soucieuse d'harmoniser les pratiques du recensement des populations à l'échelle mondiale, se soit chargée de fournir des recommandations pour la définition des périmètres urbains (Nations Unies, 1998, p. 68, 2009, p. 132 ; Nations Unies et EUROSTAT, 2006, p. 45). Le critère qu'elle valorise alors est celui de la continuité du bâti : « Selon les recommandations de l'ONU, la distance entre deux constructions permet de définir les limites d'une agglomération. Elle ne doit pas excéder 200

mètres en Europe, 500 mètres en Amérique du Sud où le système de peuplement est plus lâche. Au-delà de cette distance, la continuité du bâti est rompue : on sort des limites "morphologiques" de l'agglomération. » (Moriconi-Ebrard, *op. cit.*).

En France, cette recommandation est celle suivie par l'INSEE pour délimiter les « unités urbaines » (Julien, 2000). En complément, l'INSEE se charge de réaliser un zonage par « aires urbaines » qui tient compte de la fréquence des déplacements domicile-travail enregistrés entre les unités urbaines et leur périphérie (Brutel & Levy, 2011 ; Déjoie-Larnaudie, 2011). Puisque l'ONU a très tôt mis en avant le principe d'une définition morphologique ou physique de la ville à l'échelle mondiale, Moriconi-Ebrard l'a repris pour concevoir les agglomérations de la base GÉOPOLIS. Par souci de comparabilité et d'accessibilité des données à l'échelle mondiale, il a renoncé à utiliser une méthode de délimitation s'appuyant sur un trop grand nombre de critères. Dans la mesure du possible, c'est-à-dire dans les cas où ils étaient comparables, il a réutilisé des périmètres existants, mais quand la définition officielle du pays s'appuyait sur un critère fonctionnel, il a redéfini la limite des agglomérations en appliquant uniquement le critère morphologique (Figure 24).

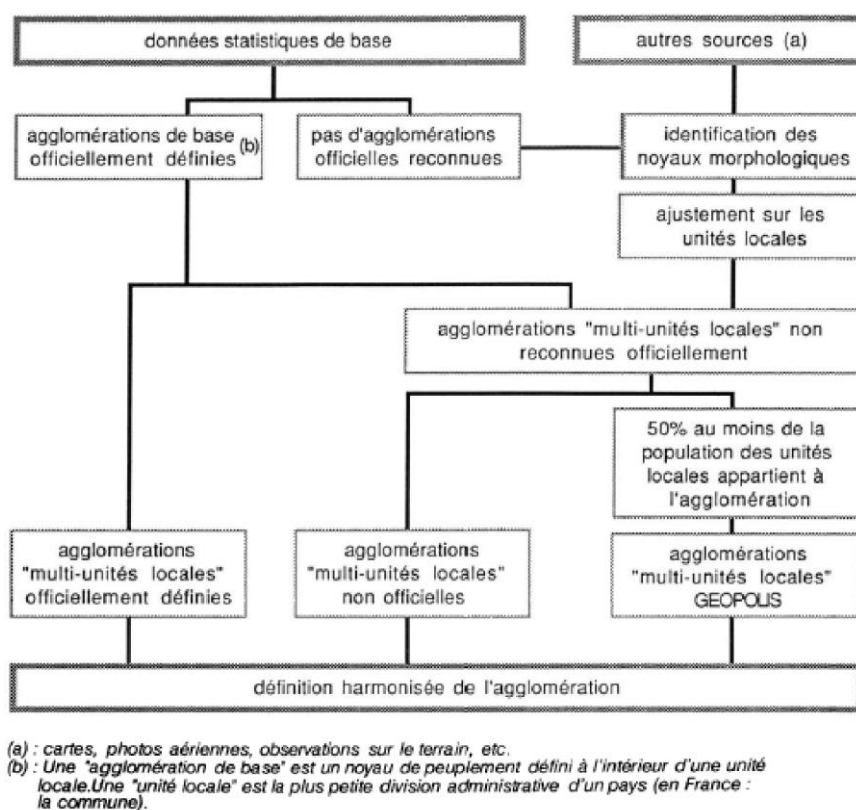


Figure 24 – Les étapes d'une définition harmonisée de l'agglomération. Schéma extrait de l'article « Les 100 plus grandes villes du monde » (Moriconi-Ebrard, 1991)

L'avantage de son approche est que la morphologie d'une ville est plus stable que son « fonctionnement » mesuré à partir de données beaucoup plus fluctuantes comme les déplacements domicile-travail. Car l'enjeu de la comparabilité des données n'est pas seulement spatial, il est aussi temporel (Julien, 2000). Toutefois, même si elles évoluent moins vite que les pratiques urbaines, les limites morphologiques des villes sont amenées à progresser avec le temps, ce qui justifie des révisions périodiques de la base GÉOPOLIS. Au cours des années 2000, les données d'observation spatiale librement accessibles sur *Google Earth* ont ainsi été mobilisées pour mettre à jour le projet e-Géopolis⁵⁴.

Alors que le recueil de données objectives sur la morphologie des villes était déjà plus réaliste à mettre en œuvre dans les années 1990 que le recueil de données fonctionnelles sur chaque système urbain à l'échelle mondiale, il s'est encore simplifié au cours des années 2000. En effet, plusieurs groupes de chercheurs ont travaillé à mettre au point des méthodes automatisées et de moins en moins fastidieuses pour détecter des taches urbaines à partir de données satellitaires. L'article de Johanna Baro *et al.* paru dans *Cybergeo* en 2014 rend compte de plusieurs avancées en ce domaine (Baro, Mering, & Vachier, 2014). Plus généralement, les progrès en télédétection et l'amélioration de la résolution des images satellitaires depuis le début des années 2000 ont permis d'envisager des études fines de l'occupation du sol à l'échelle planétaire⁵⁵ ; tandis que les capacités de traitements automatiques ont augmenté grâce au perfectionnement des Systèmes d'Information Géographique (SIG) assistés par ordinateur. Ainsi, entre autres travaux, Annemarie Schneider *et al.* décrivent en 2010 une nouvelle base d'utilisation du sol construite à partir des données des instruments d'observation satellitaires MODIS à 500 mètres de résolution et l'application d'une méthode de détection automatique d'aires urbaines testée sur un échantillon de 140 villes (Schneider, Friedl, & Potere, 2010). En plus de proposer une définition harmonisée du fait urbain, ce dernier article incite à se méfier des bases de données mondiales sur l'emprise urbaine parce qu'elles s'appuient sur des informations issues de différents satellites, pouvant faire aboutir à différents résultats en matière de continuité du bâti. Face à cette difficulté, une spécialité de recherche s'est développée qui envisage l'extraction et la conception d'objets géographiques à partir d'images à haute résolution. Cette approche appelée GEOBIA pour *Geographic object-based image analysis* a déjà fait l'objet de plusieurs conférences et d'un numéro spécial de la revue *International Journal of Applied Earth Observation and Geoinformation* en 2012 (Addink, Van Coillie, & De Jong, 2012).

⁵⁴ L'objectif du projet e-Géopolis est de mettre en ligne et à jour la base GÉOPOLIS.

⁵⁵ Qu'il me soit permis de remercier ici Pauline Crombette, ma « co-bureau », de m'avoir sensibilisée à cette question qui se trouve au cœur de sa thèse.

En contrepoint, bien qu'ils soient plus difficiles à recueillir à l'échelle mondiale et moins stables dans le temps, les critères fonctionnels de définition des espaces urbains sont, à certains égards, considérés comme plus pertinents que les critères morphologiques.

Au cours des années 2000, certains chercheurs estiment même les préconisations de l'ONU obsolètes pour définir l'urbain : « Cette définition répandue par les Nations Unies s'appuie sur une vision anachronique des unités urbaines, en particulier en ce qui concerne les études comparatives, parce que la distance entre les bâtiments ne détermine en rien la fonction, même si la distance influence la fonction. »⁵⁶ (Matthiessen & Schwarz, 1999). Depuis les années 1950, la référence en matière d'aire fonctionnelle est incarnée par les aires métropolitaines américaines ou *Standard Metropolitan Areas* (SMA) établies lors du recensement américain. Ces aires fonctionnelles se définissent à partir des rapports entre un centre et sa périphérie, la périphérie obéissant à des exigences en termes de densité de population et de structure d'activité. À la fin des années 1960, un nouvel ensemble d'aires urbaines américaines est défini sous la pression de Brian Berry : les *Daily Urban Systems* (DUS) qui vient simplement s'ajouter au premier ensemble sans le remplacer. Il s'agit d'aires urbaines élargies dont « le centre correspond à une SMA dans les aires fortement urbanisées, ou à un comté dont la population est de 25 000 à 50 000 habitants dans les parties rurales » (Cattan & Rozenblat, *op. cit.* ; Berry, Goheen, & Goldstein, 1969, p. 32). Cette fois-ci les unités formant la périphérie (comtés contigus) sont sélectionnées en tenant compte de la fréquence des navettes, de la structure et du trafic sur le réseau routier, des échanges téléphoniques, et de la diffusion des journaux. Cette définition opérationnelle aux États-Unis a été appliquée pour la première fois en Europe par des chercheurs britanniques (Cheshire, 1997). Mais des *Metropolitan Economic Labour Areas* (MELA) des années 1960 aux *Functional Urban Regions* (FUR) des années 1980, les tentatives de transposition se sont, d'après Cattan et Rozenblat, révélées peu fructueuses ou opérantes (Cattan & Rozenblat, *op. cit.*).

À la fin des années 1980, les géographes français travaillent eux aussi à définir des aires urbaines fonctionnelles à l'échelle européenne. Deux initiatives se distinguent : d'une part le travail du GIP RECLUS pour délimiter les grandes villes européennes au cas par cas à l'aide de spécialistes et, d'autre part, la volonté dans le cadre du réseau *Network for Urban Research in European Community* (NUREC) de « se donner des moyens d'harmonisation plus souples et plus adaptés à la variété du fait urbain dans les pays européens » en retenant plusieurs définitions (Cattan & Rozenblat, *ibid.*).

⁵⁶ L'original : « This widely used UN definition is actually an anachronistic delimitation of greater urban units, especially when it comes to comparative studies, because the distance between buildings does not determine function, although distances influence function. » (Matthiessen & Schwartz, 1999).

Au cours des années 2000, à l'occasion d'un programme de recherche européen sur les métropoles compétitives sous la direction de Michaela Paal appelé COMET, les périmètres urbains obtenus par le NUREC ont été mis à jour à l'aide d'un Système d'Information Géographique (Paal, 2006). Par ailleurs, bénéficiant du soutien de la Commission Européenne et d'Eurostat, les efforts pour obtenir une définition harmonisée des aires fonctionnelles à l'échelle européenne se sont poursuivis. Au début des années 2000, deux nouveaux projets se mettent en place : celui du GEMACA et celui de l'Audit Urbain.

Le GEMACA pour *Group for European Metropolitan Areas Comparative Analysis* est porté par des chercheurs britanniques à partir de la fin des années 1990. 14 régions urbaines sont définies par le GEMACA dans le cadre du programme européen de coopération interrégionale pour les régions métropolitaines du Nord-Ouest de l'Europe (INTERREG IIC). Celle de Londres, notamment, a été définie par Alan Freeman (A. Freeman, 2007a, 2007b). Deux particularités font l'originalité de ce projet et expliquent qu'il ne soit pas facilement transposable : les périmètres urbains définis par le GEMACA appartiennent uniquement au Nord-Ouest européen et la définition retenue présente un caractère purement économique : le noyau de l'aire urbaine est défini à partir d'un critère de densité d'emplois plutôt qu'à l'aide d'un critère classique de densité de population ou de continuité du bâti (Comin, 2009, p. 197).

Parallèlement, Eurostat se lance dans l'Audit Urbain en 1997 (EUROSTAT, 2010). Avec la coopération de plusieurs offices nationaux de statistiques, ce projet a été l'occasion de recueillir des informations statistiques à plusieurs niveaux de résolution urbaine concernant une sélection de moyennes et grandes villes européennes. Mais, contrairement aux aires urbaines définies par le GEMACA, celles de l'Audit Urbain s'appuient sur des délimitations urbaines dont la définition varie en fonction des pays. À ce propos, Comin remarque : « la définition de la région urbaine employée par Eurostat pour l'Audit Urbain est opérationnelle pour la collecte de données statistiques ainsi que pour l'évaluation et la mise en œuvre de politiques publiques, puisque le choix des unités spatiales composant les aires urbaines fonctionnelles a été effectué dans cette perspective. Cependant, la comparabilité internationale des villes ainsi définie n'est pas totalement assurée, puisque la délimitation des régions urbaines fonctionnelles n'est pas homogène. ». D'après elle, « on retrouve le même écueil dans les principaux travaux de comparaison internationale des performances économiques et sociales des aires urbaines de différents pays » (Comin, 2009, p. 196), c'est-à-dire dans les rapports réalisés au milieu des années 2000 par la division de l'habitat des Nations Unies (ONU-Habitat), l'OCDE (villes compétitives) ou encore Nordregio qui sont tous concernés par ce problème.

Plus près de nous et pour dépasser ces différentes limites, un collectif de chercheurs européens s'est mis à travailler, dans le cadre du programme ESPON 2000-2006, à une définition harmonisée au niveau européen des aires urbaines fonctionnelles (FUA).

Les périmètres obtenus par ce groupe en 2010-2011 sont intégrés dans la base de données ESPON en 2012, qui permet d'accéder à un grand nombre d'indicateurs statistiques à l'échelle européenne.

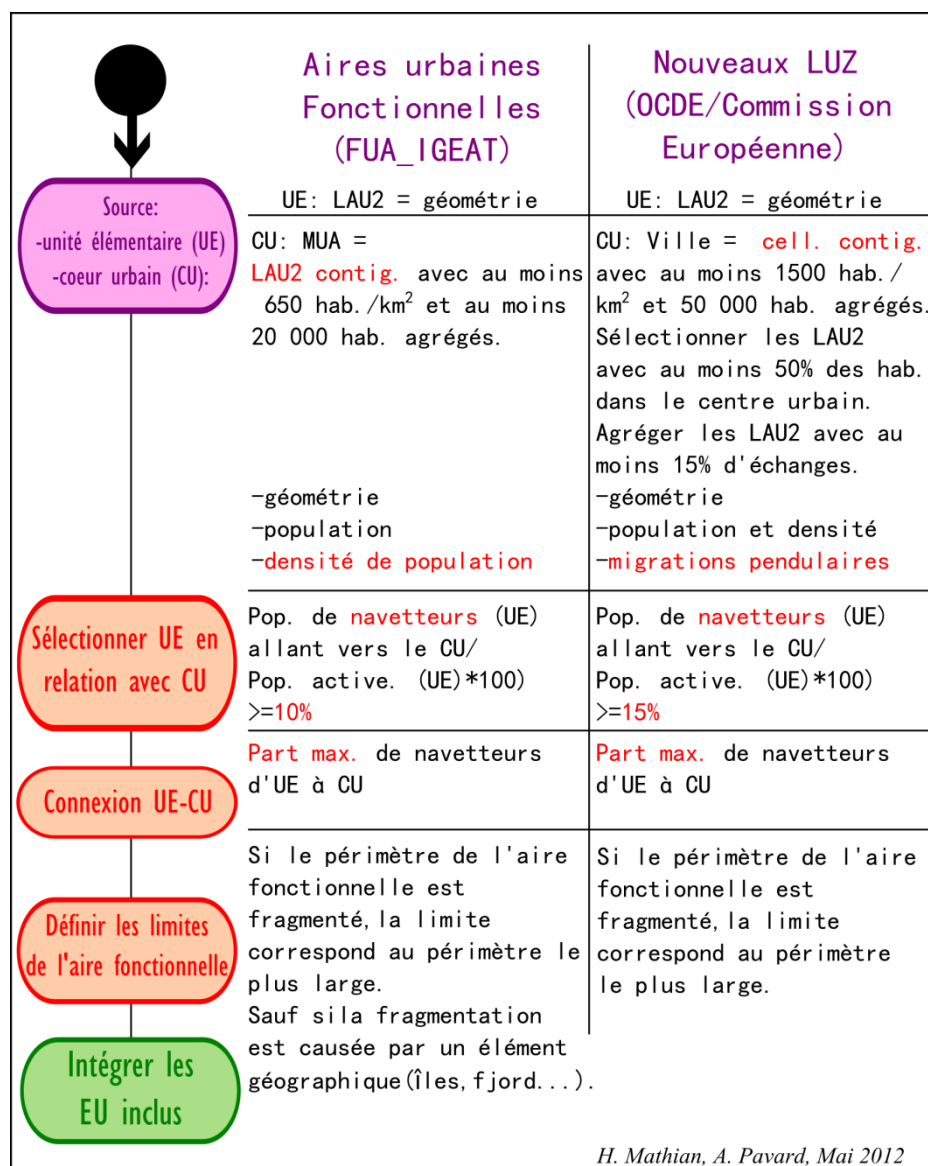


Figure 25 – Les différents processus pour implémenter les modèles conceptuels d'aires urbaines fonctionnelles. Schéma traduit et adapté du rapport technique « Functional Urban Areas (FUA) and European harmonization » (Guérois, Bretagnolle, Mathian, & Pavard, 2014)

La définition de ces aires est disponible en ligne dans le dictionnaire des unités spatiales de la base ESPON⁵⁷ (UMS RIATE et UMR Géographie-cités, 2014), où elles portent le nom de « FUA_IGEAT » pour le nom du laboratoire qui en est à l'origine (l'Institut de Gestion de l'Environnement et d'Aménagement du Territoire de l'Université Libre de Bruxelles). Cette définition est proche de celle récemment harmonisée par l'Audit Urbain des Larger Urban Zones (LUZ), qui sont également intégrées à la base de données ESPON. La Figure 25 permet de comparer les deux définitions.

Que ce soit pour les FUA_IGEAT ou les LUZ, le noyau de l'aire urbaine est sélectionné en appliquant un critère de quantité et de densité de population tandis que l'espace périphérique est délimité en appliquant un seuil à la fréquence des déplacements domicile-travail. En fait, en même temps qu'étaient mises aux points les FUA, les responsables de l'Audit Urbain se sont chargés d'améliorer leur définition des LUZ. Intégrant les critiques reçues au milieu des années 2000, l'Audit Urbain est arrivé à une version révisée des LUZ à la fin des années 2000. D'après Anne Bretagnolle *et al.*, les versions antérieures étaient fondées sur un ensemble de délimitations nationales qui dépendaient de différentes conceptions de l'urbain : « seulement 12 pays intégraient une définition s'appuyant sur les déplacements domicile-travail (avec des seuils variables) » (Bretagnolle *et al.*, 2011). En contrepoint, « la nouvelle base de données devrait donner accès à des résultats plus harmonisés », puisque « le protocole des LUZ harmonisés a déjà été adopté par 16 pays et 4 autres pays ont préféré d'autres définitions fondées elles aussi sur les navettes domicile-travail. »⁵⁸ (UMS RIATE et UMR Géographie-cités, 2014). Bien que les définitions des FUA_IGEAT et des LUZ harmonisées soient finalement proches, les périmètres obtenus ne sont pas substituables puisque les seuils utilisés pour construire les FUA_IGEAT sont plus contraignants ou discriminants que ceux des LUZ harmonisés (Figure 25). De ce fait, le nombre d'aires urbaines obtenues en appliquant la définition de l'IGEAT (1530) dépasse le nombre de celles obtenues avec la méthode de l'Audit Urbain (695). L'application de ces définitions dans l'espace européen se traduit par l'agrégation de communes contiguës (le niveau LAU2 dans la nomenclature NUTS).

Bien que limités à l'espace européen, ces efforts d'harmonisation communautaires ne sont pas passés inaperçus auprès des cellules statistiques des organismes internationaux.

⁵⁷ Source: URL:

http://database.espon.eu/db2/jsf/DicoSpatialUnits/DicoSpatialUnits_onehtml/index.html

⁵⁸ L'original : « In the previous version of LUZ produced by the Urban audit (LUZ 2004), the LUZ delineation was based on the collection of national delineations depending on different conceptual approaches of cities and different sources (Bretagnolle *et al.*, 2011). Only 12 countries had LUZ definitions based on commuters (with various minimal commuting thresholds). The new database should ensure better harmonized results. For the time being, the LUZ harmonized protocol has been adopted by 16 countries and 4 other countries have preferred other definitions based on commuters. » (UMS RIATE et UMR Géographie-cités, 2014).

En particulier, l'OCDE a décidé de transposer la définition des LUZ pour construire une population d'aires urbaines fonctionnelles à l'échelle de 28 pays de son propre périmètre. La procédure et les résultats obtenus à l'issue de cette opération sont décrits dans un rapport de l'OCDE paru en 2012 intitulé « *Redefining "Urban" : A New Way to Measure Metropolitan Areas* » (OCDE, 2012). 1148 aires urbaines fonctionnelles ont été délimitées à cette occasion. Pour les agglomérations européennes, les périmètres sont ceux des LUZ harmonisées, c'est-à-dire que le noyau des aires fonctionnelles est obtenu en appliquant un seuil à la densité de population au km² estimée par le *Joint Research Centre* (JRC) de l'Agence Européenne de l'Environnement (EEA) à partir des données de la base européenne Corine Land Cover. Pour délimiter des noyaux urbains dans le reste du monde, une procédure analogue est appliquée à partir des données LandScan issues du laboratoire d'Oak Ridge dans le Tennessee. Une fois que le noyau urbain est délimité, un seuil commun à tous les pays (15 %) est appliqué à la fréquence des déplacements domicile-travail entre entités spatiales pour définir le périmètre des aires fonctionnelles. Enfin, comme pour les LUZ harmonisées, la méthode prévoit de définir des aires polycentriques si jamais les échanges entre noyaux urbains dépassent également 15 % (*ibid.*). Parmi les premières exploitations statistiques s'appuyant sur ces entités, on note le travail de Paolo Veneri qui a contribué à leur conception (Veneri, 2013). Cette étude porte sur l'ajustement de la population des aires urbaines de l'OCDE à la loi de Zipf. En conclusion, il revient sur la difficulté d'étendre l'analyse dans le temps, en raison de l'évolution de la situation économique et de la dynamique des systèmes urbains. On retrouve la difficulté pressentie par Moriconi-Ebrard : il est moins compliqué d'obtenir des données fiables à plusieurs périodes sur l'occupation du sol et son évolution (critère morphologique) que sur les flux économiques (critère fonctionnel), surtout à l'échelle mondiale.

Si ces avancées sont prometteuses, le temps n'est pas encore venu où nous disposions de données fiables sur les déplacements domicile-travail à l'échelle mondiale, et encore moins sur leurs évolutions. Cependant, d'autres jeux de données sont maintenant accessibles à l'échelle mondiale dont aucun organisme statistique n'aurait imaginé pouvoir disposer auparavant. En particulier, les données mises à disposition à partir du suivi instantané de l'activité des internautes, des utilisateurs de téléphones mobiles, ou, plus globalement, les données issues de la généralisation du *badging* que l'on qualifie parfois de « traces », peuvent nous renseigner sur les flux intra urbains. En effet, il n'est plus très rare de disposer d'informations précises sur les temps de contact et la mobilité en temps réel des individus au quotidien (*Delay tolerant networks*). Le traitement des signaux d'appels générés par les téléphones mobiles permettent de suivre de façon inédite les mobilités (Fen-Chong & Pfaender, 2011 ; Järv, Ahas, Saluveer, Derudder, & Witlox, 2012 ; Louail

et al., 2014) ; tandis que, moins intrusive, l'observation du trafic automobile en temps réel est également envisagée (Thakur, Hui, & Helmy, 2013).

En attendant, au milieu des années 2010, il n'y a toujours pas de périmètres urbains fonctionnels obéissant à une définition harmonisée à l'échelle du monde entier. C'est pourquoi, comme nous allons le voir dans la section suivante, les chercheurs s'étant attelés à l'étude de la répartition spatiale des activités scientifiques au cours des années 2000 ont tous eu recours à différentes définitions du fait urbain.

2.2. Une méthode adaptée à l'étude des sciences

Dans le cadre des rares études géographiques sur l'activité scientifique réalisées jusqu'ici, Marie-Noëlle Comin a mis en œuvre une définition harmonisée des agglomérations intégrées à des programmes de recherche européens (Comin, 2009). Son objectif était d'étudier la géographie des projets soutenus par les Programmes Cadre de Recherche et de Développement Technologique (PCRD) consacrés aux NBIC (Nanotechnologies, Biotechnologies, Technologie de l'information, Science cognitive). Dans cette perspective, elle a jugé intéressant de se reposer sur la notion d'aires urbaines fonctionnelles (*ibid.*, p. 207-208). Ainsi, elle partage avec d'autres géographes la conviction que ce type de définition (intégrant généralement un pôle ou noyau et ses satellites) permet d'aboutir à des objets qui correspondent mieux à la « ville vécue, à travers les déplacements domicile-travail ou loisirs des populations s'éloignant de plus en plus de l'aire bâtie en continu, notamment pour bénéficier des prix fonciers plus bas ». D'après elle, ce choix est également justifié par la position de nombreux sites scientifiques en périphérie de pôles urbains structurants. Son dernier argument est emprunté à Paul Cheshire, l'un des premiers géographes européens à avoir utilisé les *Functional Urban Area* (FUR) dans les années 1980. Voici comment Comin traduit et retranscrit son propos : « Même si l'essence des villes peut être considérée comme étant leur relativement forte densité de population par rapport à leur environnement, la définition de la ville qui se prête le mieux à la comparaison internationale est plus large et correspond à sa délimitation fonctionnelle, c'est-à-dire que ses limites sont déterminées sur la base des relations économiques plutôt que sur les divisions historiques ou politiques et sont inclusives plutôt qu'exclusives. » (Cheshire cité par Comin, *loc. cit.*).

Comme souvent pour définir des aires fonctionnelles, la méthode appliquée par Comin comprend deux étapes. Dans un premier temps, les noyaux des aires fonctionnelles sont sélectionnés à l'aide des limites d'agglomérations morphologiques de plus de 10 000 habitants disponibles dans la base GÉOPOLIS. Puis, en raison de l'absence de « données comparables et fiables relatives à la mobilité quotidienne des populations en

Europe », l'étendue des aires fonctionnelles est calculée en appliquant depuis le centroïde de chaque noyau « un seuil de distance représentant un temps de trajet défini comme supportable pour les populations en ce qui concerne leurs déplacements quotidiens (domicile-travail) et l'accès aux infrastructures des agglomérations » (*ibid.*, p. 209). Plusieurs seuils sont alors considérés selon la taille des agglomérations mesurée par le nombre d'habitants qu'elles représentent : d'une distance de 30 km à partir du centroïde des agglomérations de 10 000 à 50 000 habitants à une distance de 70 km pour les agglomérations de plus de 100 000 habitants. L'application de cette méthode permet d'affecter la quasi-totalité des localités de l'étude – les localités qui participent à des programmes de recherche européens – à des aires fonctionnelles. D'après Comin, ce résultat confirme la pertinence d'une approche urbaine des activités scientifiques. Pour s'en assurer, elle se penche sur le cas des sept localités qui font exception car elles sont trop éloignées des noyaux de plus de 10 000 habitants pour en dépendre. Elle remarque que la plupart de ces localités publiantes doivent leur éloignement à la présence de ressources naturelles qui font l'objet de recherches spécifiques, et que certaines, bien que situées en ville, se trouvent sur des îles peuplées par moins de 10 000 habitants (*ibid.*, p. 218).

Ainsi, en appliquant une définition *a priori* du fait urbain, Comin parvient à agglomérer pratiquement l'ensemble des localités repérées par leur participation aux PCRD (Programmes Cadre de Recherche et de Développement Technologique) qu'elle étudie. L'intérêt de cette méthode, relativement simple, se trouve dans l'homogénéité de la population de villes obtenue. Adopter une procédure unifiée est aussi le souhait de Christian Matthiessen et ses collègues qui, dès les années 1990, mobilisent des données bibliométriques pour étudier la production scientifique des « *world cities* » (Matthiessen & Schwarz, 1999 ; Matthiessen, Schwarz, & Soren, 2002 ; Matthiessen, Schwarz, & Find, 2002 ; Matthiessen *et al.*, 2010). Comme Comin, ils privilégient une définition fonctionnelle du fait urbain et travaillent à partir d'une population d'unités urbaines définies *a priori* pour étudier la géographie des activités scientifiques. Dans leurs premières études, ils s'inspirent de la conception de la ville développée dans le cadre du NUREC⁵⁹. Puis, dans les analyses les plus récentes, ils procèdent manuellement à la délimitation des aires fonctionnelles en s'aidant de cartes topographiques, d'observations statistiques sur la densité de population et les navettes domicile-travail, ainsi que des nomenclatures métropolitaines américaines (SMA) et des découpages territoriaux européens (NUTS). Le choix d'une méthode manuelle apparaît acceptable pour étudier un nombre limité d'espaces urbains (quelques dizaines d'aires métropolitaines dans le cas de Matthiessen *et al.*), mais certainement pas pour procéder à l'étude complète du système mondial des localités scientifiques.

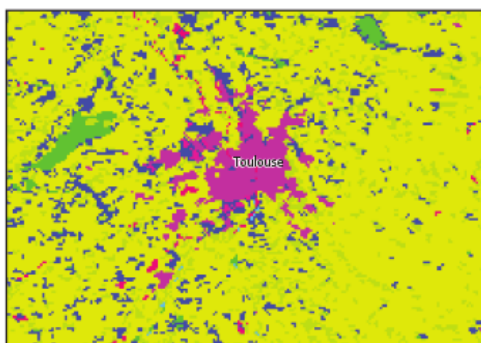
⁵⁹ Ils prévoient aussi la conception d'agglomérations polycentriques lorsque le temps de transport entre communes centres ou « noyau » est inférieur à 45 minutes (Matthiessen & Schwarz, 1999).

En effet, pour étudier la production scientifique de l'ensemble des villes du monde, il est apparu préférable d'établir un protocole et d'avoir recours à une méthode au moins partiellement automatisée.

Dans le cadre du programme « Géoscience », deux cas de figure ont été distingués. Les 500 lieux de publication les plus actifs ont bénéficié d'un traitement spécifique tandis qu'une procédure de traitement entièrement automatique a été appliquée aux localités les moins impliquées dans la production scientifique mondiale. La méthode ainsi mise en œuvre se distingue des autres dans la mesure où elle intègre la localisation des points de publications au processus de définition des agglomérations au lieu d'opter pour une définition de la ville et de s'en remettre à des périmètres urbains construits *a priori*. L'autre particularité de cette approche, c'est son ambition : il n'est pas question d'analyser uniquement la production scientifique des très grandes villes ou de se limiter à la production scientifique européenne mais de prendre en compte tous les lieux de production scientifiques du monde. Dans ces deux perspectives, l'utilisation d'un Système d'Information Géographique (SIG) est tout indiquée. D'une part, il est possible d'y reporter la position spatiale, le nom et le nombre de publications scientifiques associés aux localités publiantes préalablement géolocalisées ; d'autre part, cette information peut être croisée avec des indicateurs et des renseignements de différentes natures sur la position et la situation des lieux considérés. En l'espèce, afin de déterminer le type de périmètres urbains adapté à l'étude des activités scientifiques, le résultat du géocodage des publications a été confronté à deux définitions possibles de la ville : morphologique et démographique. Pour cela, des jeux de données mondiaux ont été mobilisés qui étaient librement accessibles en ligne et portaient sur l'occupation du sol d'une part et la densité de population d'autre part.

1) Pour l'occupation du sol, deux sources ont été envisagées : le programme *Iona GlobCover* qui utilise les instruments d'observation MERIS de l'*European Spatial Agency* (ESA) et le programme « *Global Urban Extent* » dirigé par l'équipe d'Annemarie Schneider, qui utilise les instruments d'observation MODIS de la NASA, déjà évoqués (Schneider *et al.*, 2010). D'après Eckert, Baron et Jégou ayant synthétisé la méthode : « Les données produites par ces sources présentent l'avantage d'être homogènes sur le globe et de se fonder sur des critères assez objectifs (voire automatiques) d'anthropisation des sols » (Eckert *et al.*, 2013). Pour délimiter des emprises urbaines à partir des données *Global Landcover* 2004-2006, la solution a consisté à extraire le type d'occupation du sol « *Artificial Surfaces and associated areas (urban areas > 50%)* ». Passée cette étape (voir Figure 26), les contours de l'emprise urbaine ont été automatiquement dessinés pour obtenir des formes géométriques pouvant être croisées avec les informations ponctuelles portant sur la répartition spatiale des publications scientifiques.

1- Global Landcover 2004-2006



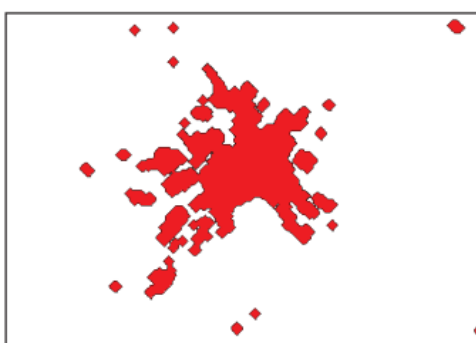
2-Extraction du type d'occupation du sol "Artificial Surfaces and associated areas (urban areas >50%)"



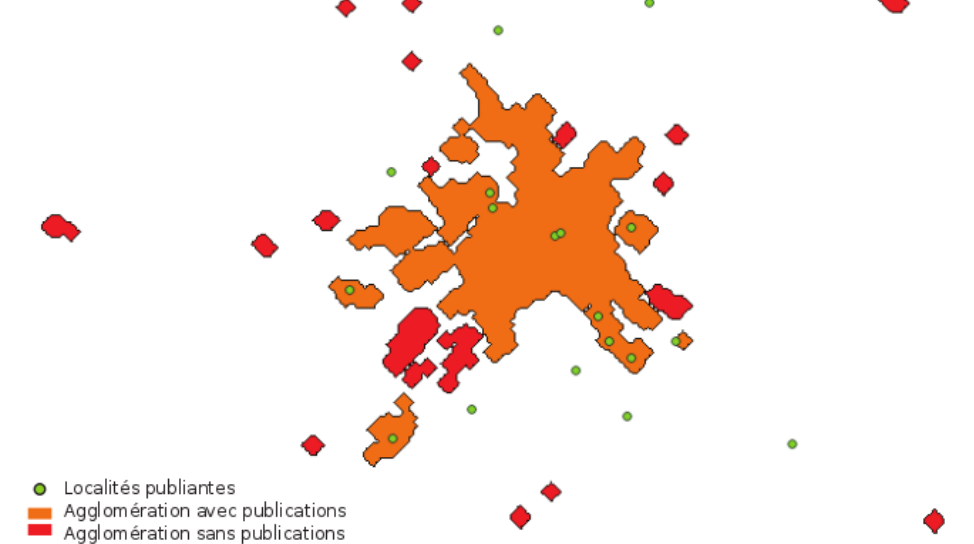
3- Lissage des contours par un filtre médian



4- Vectorisation et lissage des angles



5- Croisement avec les données de publications scientifiques (SCI Expanded)



Depuis la détection et la délimitation de "taches" urbaines jusqu'à la confrontation à la répartition spatiale des publications scientifiques: l'exemple de Toulouse par Laurent Jégou.

Figure 26 — La détection et la délimitation automatisée de « taches » urbaines au regard de la répartition spatiale des publications scientifiques : l'exemple de Toulouse par Laurent Jégou

2) Pour la densité de population : « là encore, il existe des sources de données sérieuses et de couverture mondiale, proposant des valeurs localisées de densité pour plusieurs dates. » (*ibid.*). Ce sont les données du programme *Global Population of the World*, du laboratoire SEDAC à l'Université de Columbia qui ont été retenues pour tester la pertinence du critère démographique⁶⁰ (Center for International Earth Science Information Network - CIESIN - Columbia University, United Nations Food and Agriculture Programme - FAO, et Centro Internacional de Agricultura Tropical - CIAT, 2005). En raison des fortes variations de densité à l'échelle mondiale, il n'était cependant pas possible de fixer un seuil unique valable pour l'ensemble du monde au-delà duquel il serait possible de se considérer « en ville ». Pour remédier à ce problème, il a fallu raisonner « en relatif ». La solution a consisté à employer un indicateur pour repérer les ruptures de pente dans la distribution spatiale de la densité de population, autrement dit de « déterminer spatialement les fortes variations de cette densité » (*ibid.*). En analyse spatiale, il existe des indicateurs dont c'est la fonction : les *Local Indicators of Spatial Association* (LISA), dont le I local de Moran, qui indique de façon homogène sur un territoire les regroupements statistiquement significatifs ou « noyaux de densité » (Anselin, 1995). Une fois munis des résultats obtenus sur la densité de population en appliquant cet indicateur, nous avons pu les confronter aux données de répartition spatiale des publications scientifiques dans les mêmes termes que pour l'occupation du sol.

En comparant les deux critères, celui de densité (démographique) et celui d'occupation du sol (morphologique), il est apparu que le second était moins adapté à l'étude des activités scientifiques. Le problème avec le critère morphologique est qu'il existe de nombreuses catégories de zones artificialisées et qu'il est délicat de mettre sur le même plan une zone bâtie à des fins d'exploitation touristique et une zone bâtie à d'autres fins (résidentielle par exemple). Dès lors, en appliquant le critère morphologique pour délimiter l'agglomération de Barcelone, on obtient une agglomération très étendue sur la façade maritime car intégrant l'ensemble des petites localités balnéaires qui n'ont pas grand-chose en commun avec la grande ville de Barcelone. L'agglomération ainsi constituée s'étend presque en continu le long de la côte catalane jusqu'à Tarragone (Figure 27). À l'inverse, quand on se fie aux données de densité de population, ces deux cités forment deux entités bien distinctes qu'il paraît d'autant plus pertinent de dissocier qu'il n'y a aucun lieu de production scientifique au sein du corridor les séparant comme le montre la figure suivante (Figure 29).

⁶⁰ Center for International Earth Science Information Network - CIESIN - Columbia University, United Nations Food and Agriculture Programme - FAO, et Centro Internacional de Agricultura Tropical - CIAT, « Gridded Population of the World, Version 3 (GPWv3): Population Count Grid ».

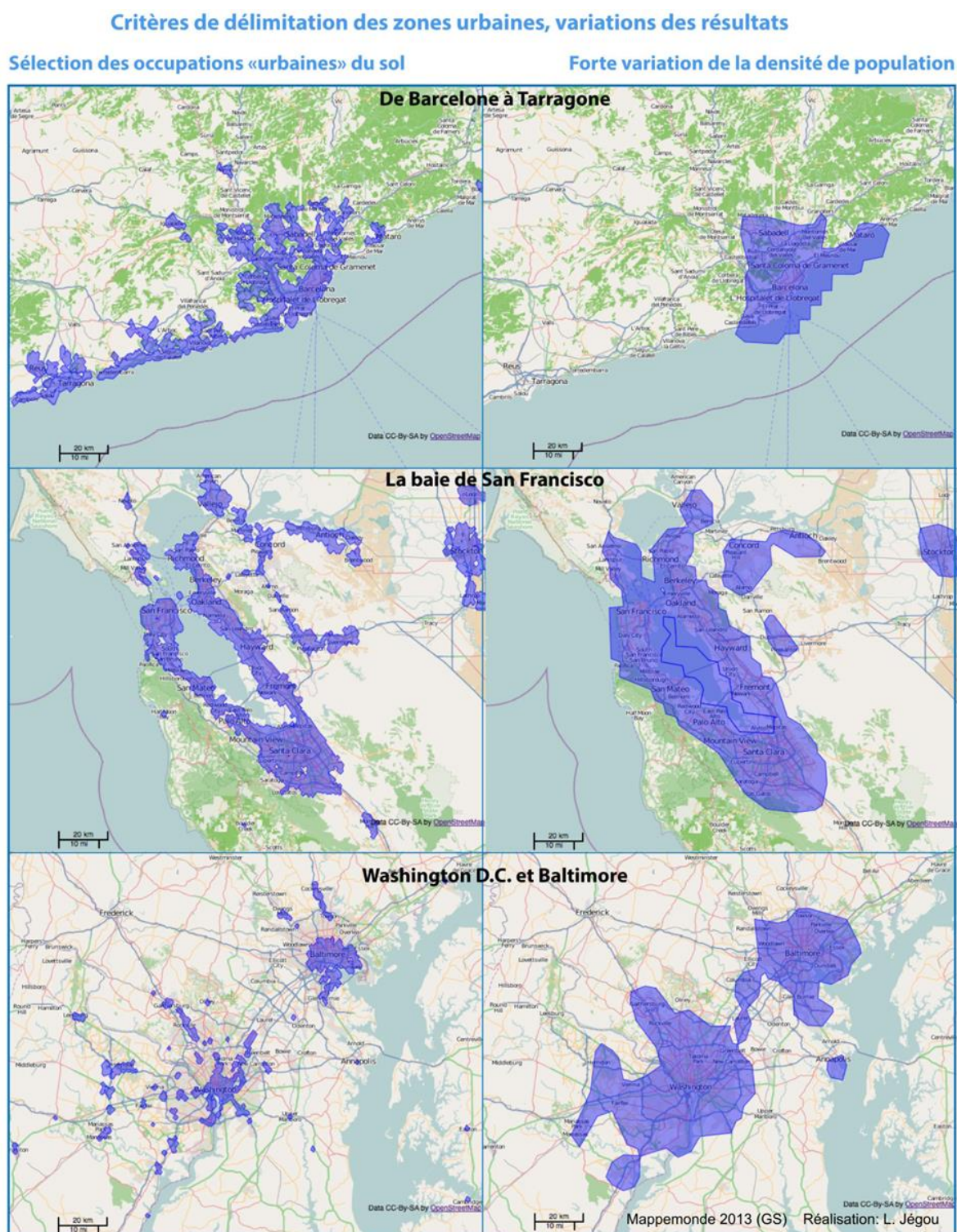


Figure 27 — Critères de délimitation des zones urbaines, variation des résultats. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)

En considérant deux options en matière de délimitation et en tenant compte de la répartition des localités publiantes à agglomérer, il est apparu que le critère de forte variation de la densité de population était le plus pertinent à mobiliser pour l'étude des activités scientifiques. Pour délimiter des agglomérations urbaines cohérentes du point de vue de la densité de population autour des 500 localités les plus actives (ayant produit le plus publié en 2008), une méthode manuelle assistée par ordinateur a été adoptée. Pour les espaces urbains denses en termes de population et de production scientifique, Denis Eckert a piloté le travail de délimitation au cas par cas, si besoin en mobilisant des avis d'experts. Les contours ainsi obtenus respectent les limites des zones de densité de population homogènes au sens d'Anselin (critère démographique), parfois combinées, en cas de doute, avec les limites des zones bâties (critère morphologique), le tout dans le respect des frontières nationales. La procédure de découpage des agglomérations a été entièrement réalisée à travers l'outil de *webmapping* interactif développé par Laurent Jégou.

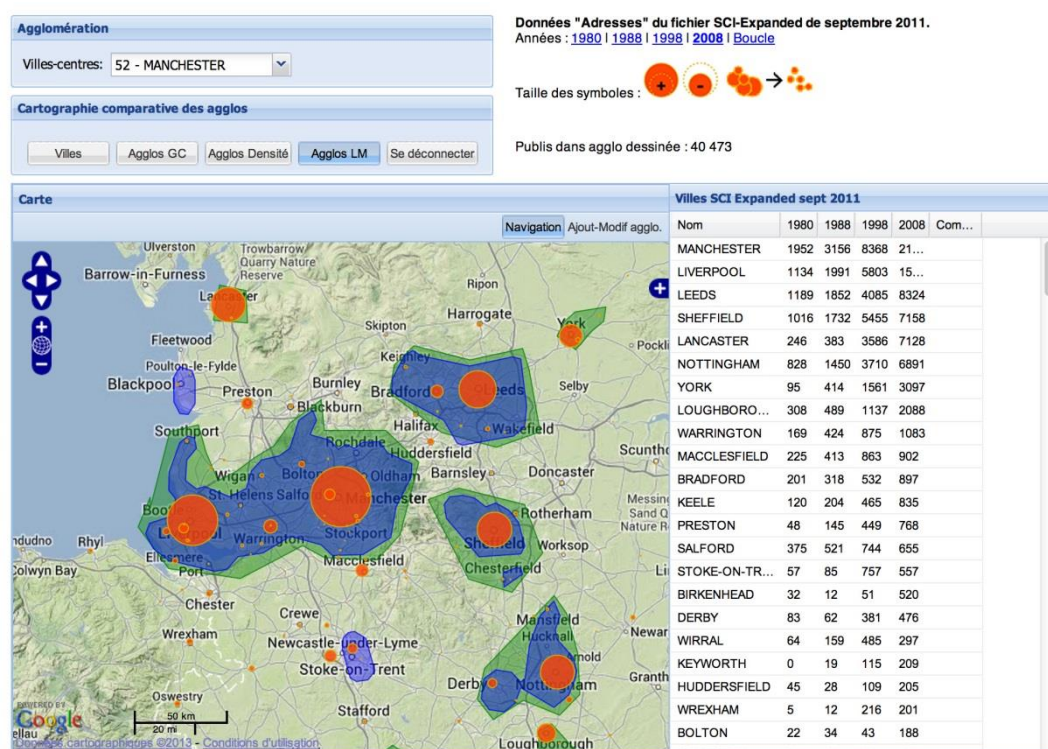


Figure 28 – L'outil d'aide à la vérification du géocodage et d'assistance à la délimitation des « agglomérations scientifiques »

La capture d'écran de l'outil (Figure 28) montre le cas de la région de Manchester-Liverpool : les agglomérations dessinées manuellement apparaissent en vert, l'emprise des zones à forte densité de population en bleu, et, en rouge-orangé, les points de publications. La taille des cercles est proportionnelle au nombre de publications.

Définition homogène des agglomérations, publications dans le WOS et hiérarchie mondiale

Publications du WOS, avant ...

... après délimitation homogène des agglomérations

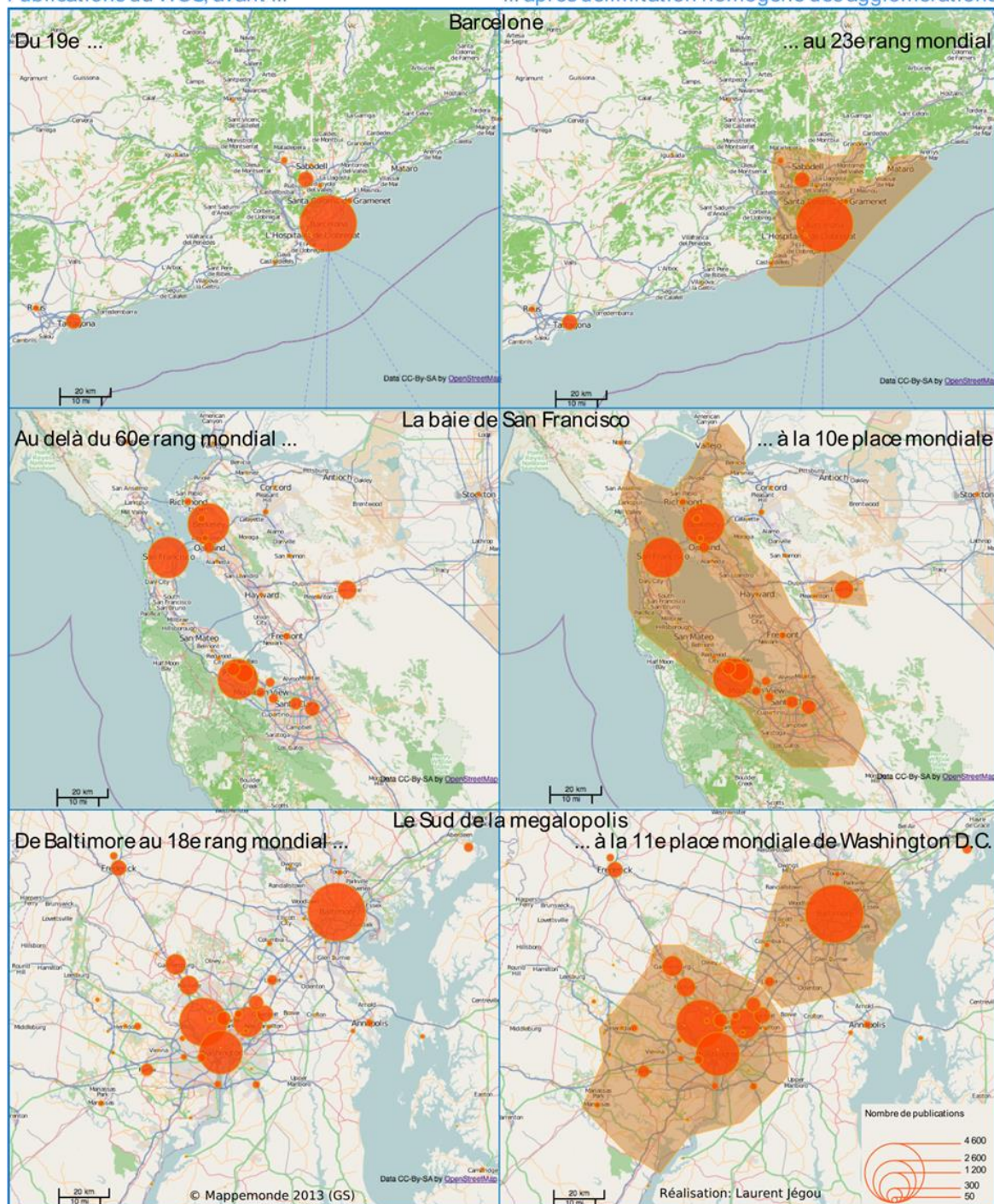


Figure 29 – Publications dans le Web of Science et hiérarchie mondiale. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)

L'intérêt d'un travail aussi minutieux pour définir l'emprise des agglomérations les plus publiantes ne fait aucun doute. En effet, cette approche a été conçue pour définir les entités spatiales qui sont susceptibles de regrouper le plus grand nombre de points de publication, et surtout les points de publications qui sont associés aux volumes de publication les plus importants. Car ce sont les unités spatiales pour lesquelles faire varier le découpage peut avoir une incidence considérable sur les résultats de l'analyse spatiale, on y reviendra. Il s'agit aussi des agglomérations qui sont le plus susceptibles d'abriter une très grande ville-centre puisque c'est dans les espaces urbains les plus denses que la densité d'activités humaines est susceptible d'être la plus importante. En 2012, cette procédure assistée a permis d'obtenir 376 agglomérations englobant les 500 premières localités (par ordre décroissant du nombre total de publications du SCI Expanded en 2008). Parmi les 500 localités les plus publiantes, certaines ont été rattachées à la même agglomération comme Manchester et Liverpool ou encore Washington et Bethesda (Figure 29), d'où le fait que le nombre d'agglomérations obtenu soit inférieur à 500.

Pour les points de publications associés à des volumes de publications moins conséquents, une procédure entièrement automatisée a, cette fois, été retenue. Les localités situées en dehors des espaces de population et de production scientifique denses ont été regroupées au sein d'agglomérations lorsque leur éloignement respectif était inférieur à 40 km. Ce critère a été appliqué en suivant un tri descendant, c'est-à-dire que les centres à partir desquels appliquer le critère de distance ont été sélectionnés dans l'ordre du plus publiant au moins publiant. Laurent Jégou a construit ces agglomérations automatiques en formulant une requête spatiale *via* l'extension PostGIS du système de gestion de base de données PostgreSQL (version 9.1). Pour chaque centre scientifique, cette requête a renvoyé la liste des localités remplissant les conditions suivantes :

- être à moins de 40 km de ce centre ;
- être associée à un nombre de publications inférieur à celui du centre ;
- ne pas déjà être agglomérée à un autre centre.

Deux types d'objets spatiaux ont ainsi été obtenus : des polygones pour les agglomérations regroupant au moins trois points de publication et des lignes pour les agglomérations regroupant seulement deux localités. Cette procédure a notamment permis de regrouper Saint-Andrew et Dundee qui sont des localités voisines trop éloignées d'Édimbourg pour être intégrées à son agglomération urbaine. Il faut être conscient qu'à l'aide de cette procédure automatique, toutes les agglomérations scientifiques possibles n'ont pas été réalisées. La procédure n'a été appliquée qu'aux localités ayant été associées à des publications en 2008 et 1998 dans le cadre du SCI Expanded. En effet, ces unités spatiales ont été conçues avant que ne soient intégrées les données obtenues auprès de l'OST Paris, notamment celles issues des autres index de citations (SSCI et AHCI).

En fait, les découpages produits et utilisés jusqu'ici sont potentiellement amenés à évoluer à l'occasion de la mise à jour qui est en œuvre actuellement, avec l'intégration des nouvelles données de publication.

En résumé, pour constituer des unités spatiales d'analyse adaptées à l'objet et aux données de production scientifique à explorer, les zones associées à d'importants volumes de publication ont fait l'objet d'un traitement spécial. Pour ces zones, « la démarche a été de privilégier le critère de l'espace "scientifique urbain" en s'intéressant à la fois au caractère urbain au sens morphologique et de la densité de population, mais aussi au critère d'une certaine densité d'activité scientifique, qui ne se superpose (spatialement) pas forcément avec le critère précédent. » (Eckert *et al.*, *op. cit.*). Pour les zones caractérisées par une plus faible densité de population et de publications, au lieu de délimiter des agglomérations urbaines, la démarche a consisté à regrouper les lieux de science partageant une proximité physique suffisamment grande pour être appréhendés à travers la même unité spatiale (le seuil de 40 km a alors été retenu). Puisque seulement les localités publiantes les plus actives ont été envisagées comme urbaines, notre définition des « agglomérations scientifiques » ne s'est pas appuyée entièrement sur une définition classique du fait urbain. Pour le dire autrement, la population d'entités spatiales ainsi obtenue ne peut être qualifiée d'urbaine que si l'on décide d'appeler « ville » tous les ensembles de proches localités dont sont issues des productions scientifiques.

En tout, 10 729 agglomérations scientifiques ont été obtenues en appliquant cette méthode aux données de publications extraites du WoS et plus précisément du SCI Expanded. En confrontant ce résultat aux données du CIESIN (Center for International Earth Science Information Network - CIESIN - Columbia University, United Nations Food and Agriculture Programme - FAO, & Centro Internacional de Agricultura Tropical – CIAT, *op. cit.*) portant sur la répartition des lieux habités (*settlements*) à l'échelle planétaire, nous avons conçu la représentation cartographique suivante (Figure 30) — obtenue en chargeant les fichiers de contours des agglomérations au format *shapefile* (fichiers de forme) (« .shp ») et le semis de localités habitées à l'aide du logiciel libre Quantum GIS (QGIS). Cette carte suggère qu'une partie non négligeable de la terre habitée accueille des lieux de publication scientifique. Bien que l'adéquation entre les deux informations soit grossière à cette échelle, elle permet d'avoir un aperçu global des espaces habités étant totalement dénués d'activité scientifique visible (vert), dont l'occurrence apparaît plus fréquente au Sud du planisphère qu'au Nord.

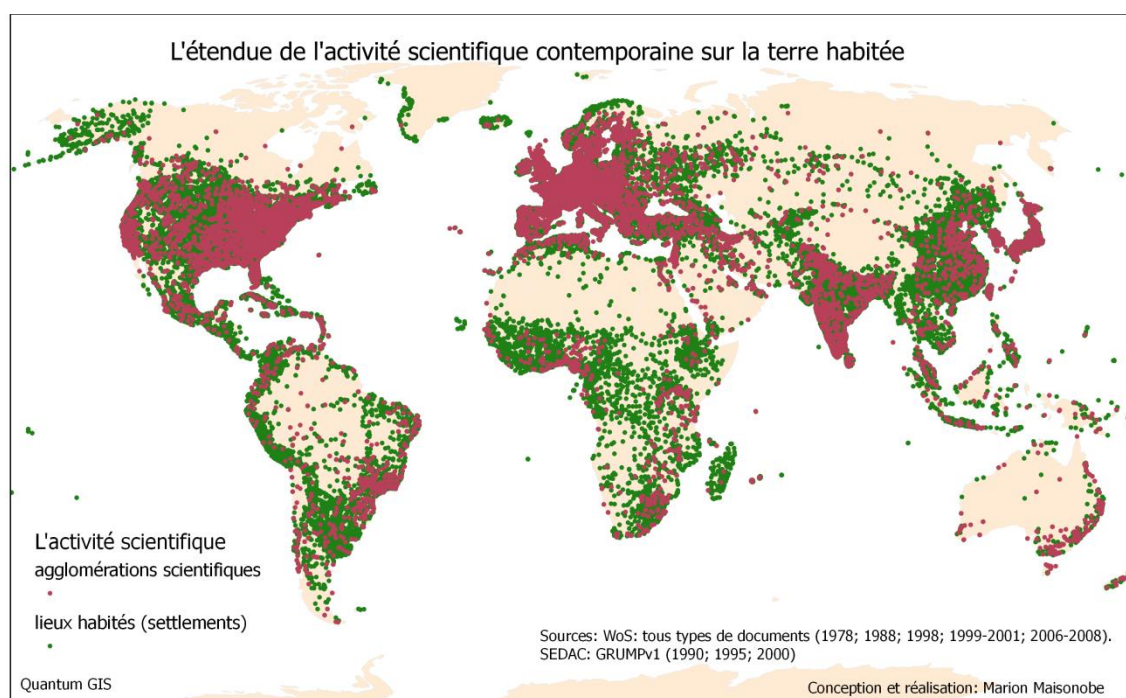


Figure 30 – Répartition des agglomérations scientifiques sur la terre habitée. Source : WoS et CIESIN

À l'issue de la procédure de géocodage des adresses de publications de l'ensemble des notices bibliographiques, il a été possible de répartir la production scientifique parmi 18 654 localités : 6059 d'entre elles ont été agglomérées manuellement au sein des 376 grandes entités spatiales tracées semi-manuellement ; 2480 ont été intégrées dans des agglomérations automatiques de type « polygone » ; 109 dans des agglomérations automatiques de type « ligne » ; et 10 006 sont restées isolées. Ces dernières agglomérations, les 10 006 localités isolées, n'ont contribué qu'à 8 % de la production scientifique totale en 2007 (moyenne mobile d'ordre 3 sur l'intégralité de la production indexée dans le WoS en 2006-2008).

Dès lors, pour produire les résultats présentés dans le Chapitre 9, la table d'appartenance géographique, obtenue par la méthode décrite plus haut, a été appliquée aux six années d'enregistrements du WoS (les 3 index confondus) recueillies auprès de l'OST Paris (1999-2001 ; 2006-2008) soit plus de 8,6 millions de publications renvoyant vers plus de 14,8 millions d'adresses. Entre 97 et 98 % de cette production ont été géolocalisés et intégrés à l'analyse. Pour arriver à de tels résultats, une étape intermédiaire s'est révélée indispensable : celle de l'agrégation des données de publication au niveau des agglomérations scientifiques. Cette étape de synthèse de l'information implique le choix d'un mode de comptage de la production, de normalisation des écarts annuels pour limiter l'impact des petites fluctuations transitoires, et la mise en forme de matrices de collaboration entre lieux de sciences à partir des données de co-signatures.

CHAPITRE 7. SYNTHÉTISER LES DONNÉES BIBLIOGRAPHIQUES

Après avoir défini les entités élémentaires d'analyse, l'étape suivante consiste à y associer des volumes de production scientifique. Il s'agit de l'opération de synthèse de l'information qui demande de faire des choix méthodologiques indispensables pour passer à l'analyse et à la représentation des résultats. Comme l'explique Denise Pumain dans son manuel d'analyse spatiale : « le plus souvent, le passage d'une échelle géographique à une autre échelle s'accompagne d'une abstraction, d'une modélisation de l'information, qui repose sur un processus d'élimination de détail appelé *généralisation*. » (Pumain & Saint-Julien, 1997, p. 37). Le développement qui suit aborde les moyens et les choix méthodologiques mis en œuvre pour produire :

- Un nombre de publications scientifiques par lieu, et par année ;
- Un nombre de relations entre ces lieux compte tenu des co-publications, par année.

Théoriquement, il est également possible de déterminer un nombre de citations reçues et un nombre de citations envoyées par lieu mais cette possibilité n'a pas encore été pleinement explorée par notre équipe.

Dans un premier temps, nous décrivons la méthode de comptage retenue pour mesurer des nombres de publications par lieu par année. À titre illustratif, nous montrons l'influence du niveau d'agrégation de l'information et l'influence du mode de comptage sur la hiérarchie des lieux de production aux Pays-Bas. Ce constat conduit à mettre en évidence les dessous d'un enjeu stratégique regardant les pratiques de signature des articles par les auteurs de publications scientifiques. Dans un deuxième temps, nous montrons comment extraire une matrice relationnelle à partir des données de publications. L'opportunité d'extraire des graphes de collaborations scientifiques à partir d'une matrice de lieux est évoquée. Nous considérons, pour finir, l'influence des périmètres d'agglomérations sur la structure des réseaux de collaboration entre agglomérations scientifiques. C'est l'occasion d'interroger le sens éventuel d'une prise en compte des relations scientifiques entre lieux pour la construction des entités spatiales d'analyse.

1. Le traitement des données de production

*« L'utilisation de mesures quantitatives représente toujours un "filtre" qui permet d'extraire une propriété caractéristique de la structure empirique. Ainsi, il s'agit d'une part d'une réduction d'information, d'autre part cette réduction sert à obtenir une notion approfondie de la structure observée qui est d'abord voilée par des particularités sans importance. Une telle méthode sélective permet donc de connaître des propriétés plus générales qui n'apparaissent pas immédiatement. »
(Frankhauser, 1994)*

1.1. Le comptage de la production scientifique

Pour associer des nombres de publications aux agglomérations scientifiques, il existe plusieurs méthodes. Privilégier une méthode sur une autre suppose de faire varier la façon dont on traite les publications écrites en collaboration ou, du reste, signées depuis plusieurs endroits (celles dont le ou les auteurs signalent au moins deux adresses). Pour ce type de publications multi-localisées, il existe plusieurs possibilités de comptage. Pour la majorité des traitements réalisés dans le cadre du programme de recherche « Géoscience », la méthode retenue est celle dite de *Whole Normalized Counting* (Gauffriau, Larsen, Maye, Roulin-Perriard, & Ins, 2008), ou comptage entier fractionné :

- *Whole*, ou « entière », parce que l'unité de compte n'est pas le nombre d'adresses par publication, mais le nombre d'entités spatiales par publication (l'entité spatiale à laquelle on se réfère pour décrire la méthode est l'agglomération scientifique préalablement définie) ;
- *Normalized*, parce qu'on attribue une valeur fractionnée à chaque agglomération ayant contribué à la publication (chaque agglomération reçoit une fraction comme crédit par publication qui est égale à 1 divisé par le nombre d'agglomérations impliquées).

Autrement dit, si pour un article donné, il y a trois adresses à Toulouse, deux adresses dans la périphérie de Toulouse, une adresse à Paris et deux dans sa périphérie (8 adresses au total) ; le choix d'un comptage entier fractionné au niveau de l'agglomération consiste à fractionner par le nombre d'agglomérations impliquées dans la publication, c'est-à-dire deux (Paris et Toulouse), si bien que l'agglomération de Paris

reçoit un crédit de 0,5 et l'agglomération de Toulouse un crédit de 0,5. L'esprit de cette méthode permet de se concentrer uniquement sur le nombre d'entités spatiales de base qui contribuent à la publication sans prêter attention au nombre d'adresses totales qui se rattachent respectivement à chaque entité spatiale. L'exemple de la publication dont la notice est montrée ci-dessous (Figure 31, obtenue à partir d'une capture d'écran depuis l'interface en ligne du *Web of Science*) correspond à cette configuration. Dans le *Web of Science* (WoS), un identifiant est attribué à chaque notice de publication, auquel est associé le titre, les auteurs, le nom de la revue, le numéro, l'année de parution, ainsi que les adresses institutionnelles des auteurs. En général, d'autres informations sont également disponibles : le type de publication (article, recension, lettre, note de recherche, éditorial ou encore acte de colloque), les mots-clefs, le résumé en anglais, les références, les citations, et l'organisme qui, le cas échéant, a financé la recherche.

Blood levels of bacterial DNA predicts the onset of diabetes in the general population

By: Amar, J (Amar, J.)^[1]; Chabo, C (Chabo, C.)^[2]; Lange, C (Lange, C.)^[3]; Serino, M (Serino, M.)^[2]; Iacovoni, J (Iacovoni, J.)^[2]; Mondo, S (Mondo, S.)^[4]; Lepage, P (Lepage, P.)^[4]; Mariette, J (Mariette, J.)^[5]; Lantieri, O (Lantieri, O.)^[6]; Perez, L (Perez, L.)^[2]; Marre, M (Marre, M.)^[6]; Klopp, P (Klopp, P.)^[2]; Courtney, M (Courtney, M.)^[7]; Charles, MA (Charles, M. A.)^[3]; Balkau, B (Balkau, B.)^[3]; Burcelin, R (Burcelin, R.)^[2]...Less

DIABETES & METABOLISM

Volume: 37 Pages: A21-A21 Special Issue: 1

Published: MAR 2011

[View Journal Information](#)

Author Information

Addresses:

- + [1] INSERM U558, **Toulouse**, France
- + [2] INSERM U858, **Toulouse**, France
- + [3] INSERM U1018, **Villejuif**, France
- [4] INRA UMR1319, Jouy En Josas, France
- [5] Plateforme Bioinformat **Toulouse** Genopole, Castanet Tolosan, France
- + [6] INSERM U695, **Paris**, France
- [7] VAIOMER SAS, **Toulouse**, France
- + [8] INRA, GENOTOUL Platform, Auzeville, France

Figure 31 — Notice de publication signée par 16 chercheurs répartis entre Toulouse et Paris.

Source : *Web of Science*

Cet exemple de publication, parue en 2011 et pour laquelle on dispose d'informations sur le lien entre auteurs et adresses (la petite note permettant d'associer les auteurs à leur adresses, dont on ne dispose pas pour les publications indexées dans le WoS avant 2008), montre que le nombre d'adresses est différent du nombre d'auteurs attachés à la publication. Ici le rapport entre le nombre d'auteurs et le nombre d'adresses reste constant au niveau des agglomérations (6 auteurs dans l'agglomération de Paris pour 3 adresses, et 10 auteurs dans l'agglomération de Toulouse, pour 5 adresses). Mais, bien

qu'il y ait en moyenne 2 auteurs par adresse, le Tableau 10 indique que ce rapport n'est pas stable au niveau des municipalités impliquées dans la publication.

Municipalité	Agglomération	Nombre d'auteurs	Nombre d'adresses
Toulouse	Toulouse	8	3
Castanet Tolosan	Toulouse	1	1
Auzeville	Toulouse	1	1
Paris	Paris	1	1
Villejuif	Paris	3	1
Jouy en Josas	Paris	2	1
Total	2	16	8

Tableau 10 – Résumé de l'information spatiale issue d'une notice de publication (l'exemple de la Figure 31)

Dès lors, il existe différentes solutions de comptage possibles qui dépendent à la fois de l'entité élémentaire d'analyse⁶¹ et du choix de l'unité de compte. Lorsque l'unité de compte diffère de l'entité élémentaire d'analyse (dans le cas, par exemple, où on compte le nombre d'adresses par agglomération), on a affaire à un comptage « complet ».

Cette distinction est proposée dans un article de Marianne Gauffriau *et al.* paru dans la revue *Scientometrics* en 2008 et intitulé « *Comparisons of results of publication counting using different methods* » (Gauffriau *et al.*, 2008). Elle et ses collègues y décrivent différentes méthodes de comptage en prenant les pays comme entité élémentaire d'analyse :

« Il s'agit d'attribuer des crédits aux entités de base – les pays – d'après la méthode de comptage retenue.

Avec le comptage complet (C) chaque pays impliqué dans la publication reçoit un crédit par adresse (autant de crédits qu'il y a d'adresses de signature ou éventuellement, si l'information est disponible, d'auteurs localisés dans ce pays) ;

Avec le comptage complet normalisé ou fractionné (CN) tous les pays partagent un seul crédit qui est divisé d'après la répartition des adresses ou bien des auteurs ;

Dans le comptage direct (S) le premier pays signataire reçoit un crédit pour lui tout seul au détriment des autres pays signataires ;

Avec le comptage entier (W) tous les pays impliqués reçoivent un crédit de 1 peu importe le nombre d'adresses ou d'auteurs associé à chacun des pays signataires.

⁶¹ Le niveau retenu pour une analyse scientométrique est à choisir parmi une grande variété de possibilités : l'auteur, l'équipe, le laboratoire, l'établissement de recherche, la commune, l'agglomération, le pays, la région du monde, ce qui multiplie les résultats possibles !

Avec le comptage entier si un pays apparaît au moins une fois dans la liste d'adresses il reçoit un crédit entier et avec le comptage entier fractionné (WN) tous les pays impliqués au moins une fois partagent un crédit de 1.

La différence majeure entre le comptage entier et le comptage complet c'est qu'avec le comptage complet, le pays reçoit un crédit chaque fois qu'il apparaît dans la liste d'adresses. » ⁶² (*ibid.*).

En reprenant l'exemple de la publication précédente (Figure 31) et en prenant l'agglomération comme entité élémentaire d'analyse, le Tableau 11 montre que la valeur associée à la participation des entités spatiales dépend de l'unité de compte. Avec la méthode que nous privilégions (comptage entier fractionné), l'unité de compte se confond avec l'entité élémentaire d'analyse (l'agglomération) de sorte que le résultat ignore le nombre d'auteurs ou d'adresses contenus dans l'entité élémentaire d'analyse — contrairement à ce qu'il se passe en cas de comptage complet. Avec un comptage complet prenant pour unité de compte les organisations, la distribution des crédits possibles serait encore différente puisque 4 organisations sont impliquées dans cette publication : l'INSERM, l'INSA, le génopôle de Toulouse et une entreprise du nom de Vaiomer SAS.

L'exemple d'une publication produite entre les agglomérations de Paris et Toulouse

Crédits par unité spatiale selon la méthode de comptage	Compte complet		Compte complet fractionné		Compte direct	Compte entier	Compte entier fractionné
Unité de compte	Auteur	Adresse	Auteur	Adresse		Agglomération	
Paris (agglomération)	4	4	4/14	4/5	1	1	0,5
Toulouse (agglomération)	10	1	10/14	1/5	0	1	0,5

Tableau 11 — Les différents moyens de mesurer la participation des entités spatiales contribuant à une même publication (l'exemple de la Figure 31)

Que l'on prenne l'entité spatiale, l'organisation ou l'adresse comme unité de compte, l'intérêt du fractionnement est qu'en faisant la somme des fractions, on retrouve le nombre total de publication. Le fractionnement permet de traiter la variable « nombre de publications » comme un effectif. D'après Denise Pumain, les effectifs sont des « attributs additifs » : « Il suffit de faire la somme des valeurs associées à chacune des entités spa-

⁶² L'original : « Credits are attributed to the basic units – countries in the present study – according to the counting method chosen. In complete counting (C) all countries get 1 credit, in complete-normalized counting (CN) all the countries in a publication share 1 credit, in straight counting (S) the first country gets 1 credit, in whole counting (W) all unique countries will receive a credit of 1. In whole counting if a country appears more than once in the address list it will still get only 1 credit and in whole-normalized counting (WN) all the unique countries share 1 credit. In complete counting the country gets 1 credit each time it is appearing in the address list. » (Gauffriau *et al.*, 2008)

tiales de niveau inférieur pour trouver la valeur que l'on associera à l'entité de niveau supérieur qui les regroupe. » (Pumain & Saint-Julien, 1997, p. 41). En fractionnant les données de production au niveau des agglomérations, il est ensuite possible à l'aide de simples sommes d'obtenir des scores de production au niveau des pays sans doubles comptes.

En plus de montrer la variété des options possibles, l'article de Gauffriau *et al.* permet d'attirer l'attention sur le fait que préciser l'entité élémentaire d'analyse, et, au cas où elle diffère, l'unité de compte retenue pour mesurer des volumes de publications, est important car, selon les choix qui sont faits, les résultats seront différents, en particulier pour les entités d'analyse situées en milieu de hiérarchie, c'est-à-dire entre les espaces qui sont de loin les plus actifs et ceux qui ont une participation très ponctuelle à la production scientifique.

1.2. L'exemple de la production néerlandaise

C'est ce que nous avons pu observer en détail en nous penchant sur le cas particulier de la production scientifique néerlandaise. Les Pays-Bas sont un pays sur lequel il est utile de s'attarder pour au moins deux raisons :

- d'une part, c'est un pays où la densité de population est très élevée ce qui rend particulièrement délicate l'opération de découpage d'agglomérations scientifiques sur la base d'un critère démographique. Il pourrait d'ailleurs être intéressant, même si cette intention dépasse le cadre de la présente thèse, d'étudier la pertinence des périmètres délimités par notre équipe à la lueur de la structure relationnelle des centres scientifiques de ce pays ;
- d'autre part, c'est un pays particulièrement actif dans le domaine de recherche en biologie moléculaire dont nous étudions la géographie dans la troisième partie de la thèse, celui de la réparation de l'ADN.

Puisqu'elle ne s'étend qu'aux Pays-Bas, l'analyse ci-dessous a été effectuée en prenant les municipalités comme entités spatiales élémentaires. Quand il n'est pas question d'obtenir des résultats à l'échelle mondiale mais à une échelle plus fine pour les besoins d'une étude portant sur une zone géographique précise, il peut être bienvenu de changer de niveau d'agrégation des données et de privilégier un meilleur niveau de résolution que celui de l'agglomération. Rappelons que les « agglomérations » ont d'abord été conçues pour que les données de publications scientifiques soient comparables d'une zone du monde à une autre. S'il s'agit juste de comparer des données de publications à l'échelle d'un seul pays, il est légitime d'utiliser d'autres périmètres : par exemple l'ensemble des

municipalités. Pour notre exemple sur les Pays-Bas, la production scientifique du pays est agrégée au niveau municipal (Figure 32).

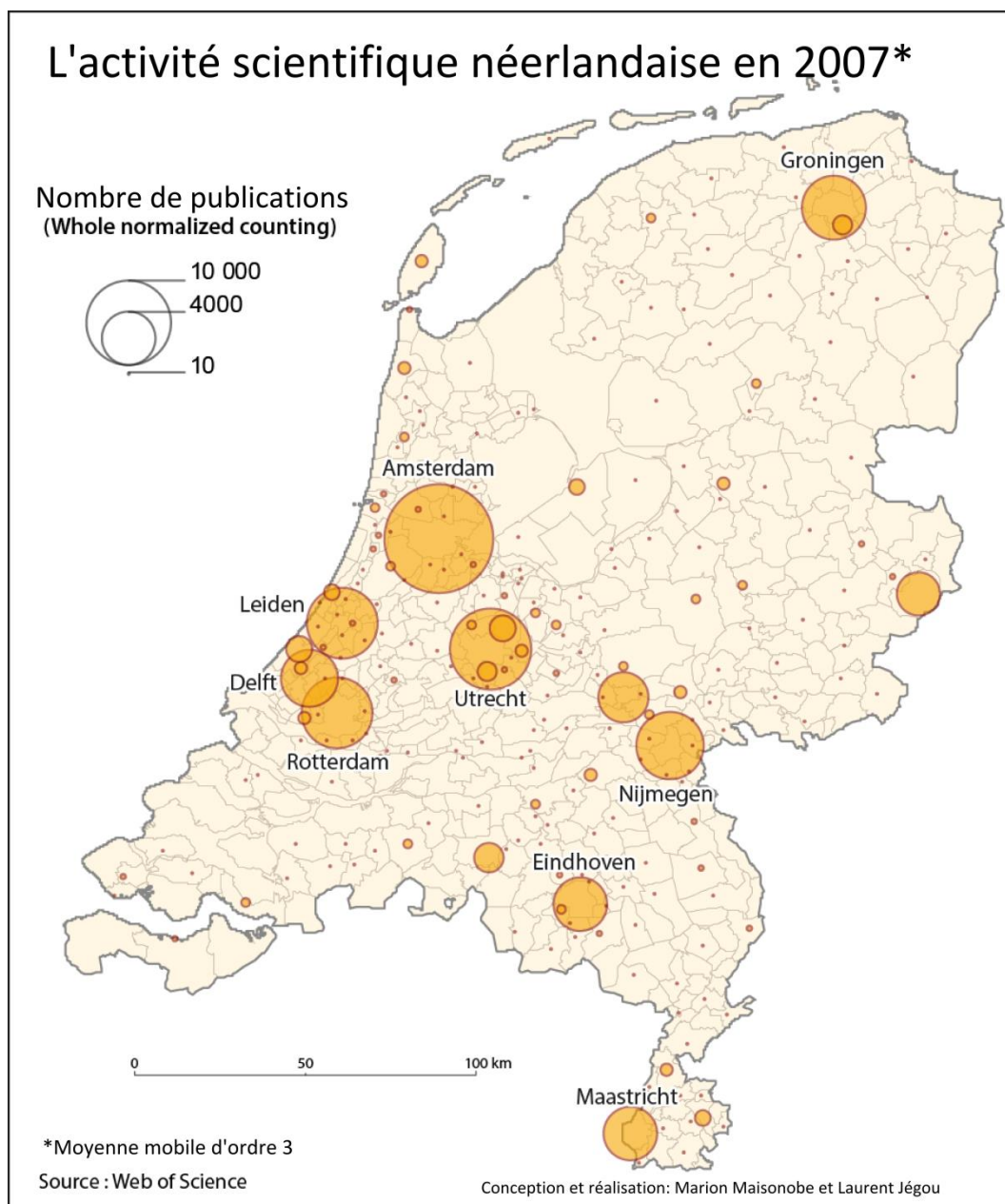


Figure 32 – Carte de la production scientifique néerlandaise en 2007 par municipalité (309 municipalités publiantes). Source : Web of Science, moyenne mobile sur 3 ans

À titre indicatif, on se propose de comparer la hiérarchie ainsi obtenue avec celle qui ressort en utilisant comme entités élémentaire les « agglomérations scientifiques ». Ce découpage est consultable sur la Figure 33 extraite de l'outil de géovisualisation.

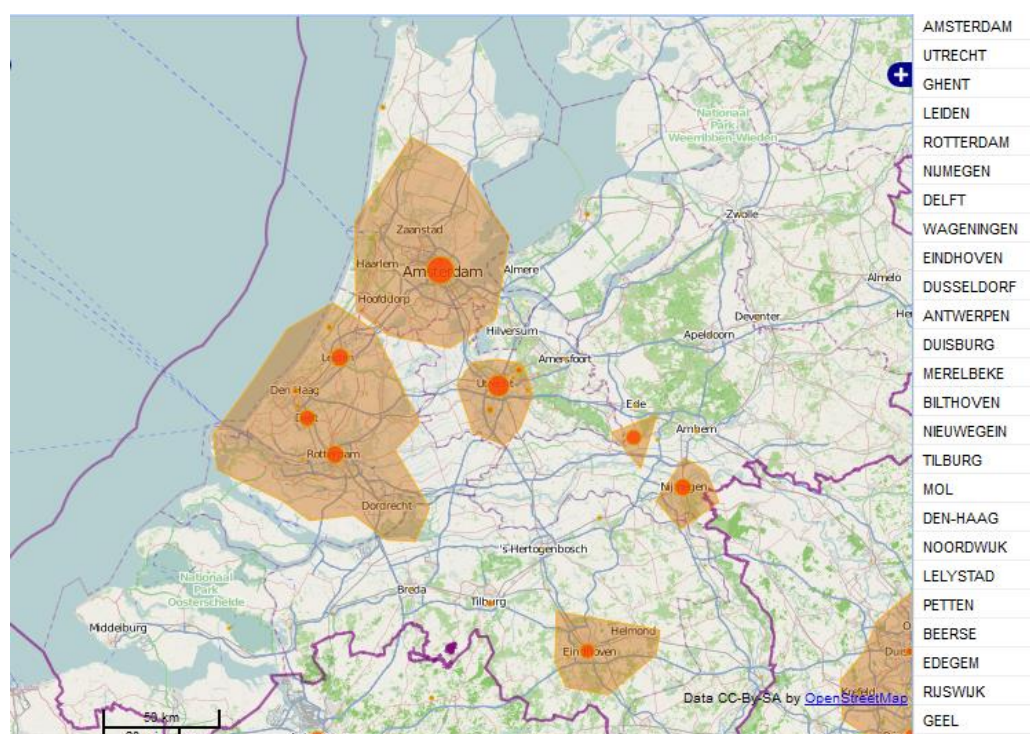


Figure 33 – La délimitation des « agglomérations scientifiques » dans la Randstad. Capture d'écran de l'outil de géovisualisation

Le Tableau 12 montre la différence dans le classement obtenu selon que l'on comptabilise les publications néerlandaises par municipalité ou par agglomération.

Rang des unités spatiales	Rang d'après le nombre de publications en 2007*	
	Municipalités	Agglomérations scientifiques
1	Amsterdam	Rotterdam-Leiden-Delft-Den Haag
2	Utrecht	Amsterdam
3	Rotterdam	Utrecht
4	Leiden	Nijmegen-Wageningen
5	Nijmegen	Groningen
6	Groningen	Eindhoven-Tilburg
7	Delft	Maastricht
8	Eindhoven	Enschede
9	Maastricht	Lelystad
10	Wageningen	Petten
11	Enschede	Zwolle
12	Tilburg	Den-Burg
13	Den-Haag	Leeuwarden
14	De-Bilt	Deventer

*Comptages entiers fractionnés (WNC), moyennes mobiles d'ordre 3

Source: SCIEp (articles, revues et lettres)

Tableau 12 – Rang des municipalités versus le rang des agglomérations à partir d'un comptage entier fractionné. Source : Web of Science, moyenne mobile sur 3 ans

Dans les deux cas, comme l'indique l'astérisque (Tableau 12), les nombres considérés pour réaliser ce classement sont les valeurs de 2007 obtenues en faisant une moyenne mobile d'ordre 3, c'est-à-dire en lissant les données de production de la série temporelle 2006-2007-2008 à l'aide de la formule suivante pour x_i la valeur prise par la variable « nombre de publications par entité spatiale » l'année i :

$$\overline{x}_{2007} = \frac{(x_{2006} + x_{2007} + x_{2008})}{3}$$

L'intérêt de procéder à ce lissage est d'envisager des valeurs insensibles aux petites fluctuations annuelles caractéristiques du rythme de la recherche et de la diversité des pratiques des éditeurs de revues scientifiques. Par la suite, on procèdera de la même manière pour la série temporelle 1999-2000-2001 et pour la variable nombre de collaborations par couple d'entités spatiales.

L'ensemble de notices bibliographiques traité dans cet exemple comprend l'intégralité des notices indexées dans le WoS lors des trois années consécutives 2006-2007-2008. Cela signifie que sont intégrés aussi bien des articles que des recensions (articles de synthèse), des lettres ou encore des actes de colloque, bref tous les types de documents indexés par l'ISI Thomson Reuters — à l'exception des ouvrages qui sont indexés séparément et depuis moins longtemps.

Sur le Tableau 12, on observe que le choix de regrouper Leyde, Rotterdam, Delft et La Haye a un effet significatif puisque l'entité obtenue est celle où l'on comptabilise le plus de publications en 2007. Il y a donc eu plus de publications signées dans cette zone que dans la zone urbaine d'Amsterdam et celle d'Utrecht. Au contraire, en restant au niveau des municipalités, c'est Amsterdam suivie d'Utrecht qui sont les municipalités d'où proviennent le plus de publications en 2007. Ces subtilités sont particulièrement importantes à avoir en tête au moment de l'interprétation ou de la conception de classements d'organisations ou de villes.

Pour analyser la production nationale par municipalité, nous avons ensuite considéré plusieurs méthodes de comptage (Tableau 13). Cette expérience permet de comprendre l'impact de ce choix sur les résultats obtenus en scientométrie spatiale. Comme le suggèrent Gauffriau *et al.*, plus il y a d'auteurs et d'adresses par publication, plus ce choix est déterminant : « Il y a de grandes différences entre les scores obtenus par les différentes méthodes. Dans la littérature, il n'y a pas suffisamment d'information donnée sur les méthodes utilisées, et aucun signe d'une terminologie et d'un consensus clair sur les propriétés et les résultats des différentes méthodes. Dans les faits, la méthode entière est favo-

nable à certains pays, en particulier les pays ayant un niveau élevé de coopération internationale. Ces problèmes augmentent avec le temps parce que les coopérations nationales et internationales ne font qu'augmenter ainsi que le nombre d'auteurs par publication. Le besoin d'un effort commun pour rectifier cette situation est important. »⁶³ (*ibid.*).

Le tableau ci-dessous fait figurer la hiérarchie des municipalités néerlandaises obtenue en faisant varier le mode de comptage des publications scientifiques de ce pays pour l'année 2007 (Tableau 13).

Rang des municipalités en fonction de la méthode de comptage appliquée au nombre de publications néerlandaises produites en 2007*

Rang selon la méthode de compte	Compte non-fractionné		Compte fractionné	
	Entier	Complet	Entier	Complet
	WC	CC	WNC	CNC
1	Amsterdam			
2	Utrecht			
3	Rotterdam			
4	Leiden			
5	Nijmegen			
6	Groningen			
7	Maastricht		Delft	
8	Delft	Eindhoven		
9	Eindhoven	Delft	Maastricht	
10	Wageningen			

*moyennes mobiles d'ordre 3

Tableau 13 – Rang des municipalités en fonction de la méthode de comptage. Source : Web of Science, moyenne mobile sur 3 ans

Pour le haut de la hiérarchie, jusqu'à la 6^{ème} place, il n'y a pas de différence dans le rang des municipalités. Ensuite, selon que l'on fractionne les participations ou non, le rang de Maastricht est le 7^{ème} ou le 9^{ème}. Si le fractionnement est défavorable à Maastricht cela

⁶³ L'original : « There are big differences between scores obtained by different methods. In one instance the reduction in scores going from whole to complete-normalized (fractional) counting is 72 per cent. In the literature there is often not enough information given about methods used, and no sign of a clear and consistent terminology and of agreement on properties of and results from different methods. As a matter of fact, Whole counting is favorable to certain countries, especially countries with a high level of international cooperation. The problems are increasing with time because of the ever-increasing national and international cooperation in research and the increasing average number of authors per publication. The need for a common understanding and a joint effort to rectify the situation is stressed. » (Gauffriau *et al.*, 2008).

vient du fait que Maastricht, davantage que Delft ou Eindhoven, publie avec de nombreux partenaires à la fois. Le fractionnement fait diminuer l'impact des publications réalisées en collaboration avec un très grand nombre de partenaires tandis qu'il valorise le poids des publications issues d'une seule ville ou de quelques villes à la fois (avec un comptage entier : 0,5 pour 2 villes, 0,3 pour 3 villes...).

D'un espace à l'autre, le degré de fragmentation institutionnelle locale ou les spécificités disciplinaires peuvent expliquer les différences obtenues entre comptage entier et complet. Plus il y a d'institutions scientifiques différentes dans la ville, plus le score de la ville obtenu à partir d'un comptage complet des publications (nombre total d'adresses) sera élevé comparé au score de la ville avec un comptage entier (nombre d'articles avec au moins une adresse dans la ville). Dans cet ensemble, Eindhoven et Delft sont les deux seules villes néerlandaises dotées d'une université technologique. Elles sont donc très comparables en termes de profil disciplinaire. Dès lors, si le comptage complet bénéficie davantage à Eindhoven qu'à Delft c'est probablement parce qu'il y a une plus grande multiplicité d'affiliations institutionnelles possibles à Eindhoven qu'il n'y en a à Delft. Cela peut venir du fait qu'à Eindhoven, contrairement à Delft, se trouve une grosse entreprise collaborant beaucoup avec l'université (Philips), alors que l'Université de Delft ne dispose pas d'un tel collaborateur dans son périmètre local.

En dehors de la sensibilité au contexte institutionnel local et à sa fragmentation, les pratiques de signature varient d'une discipline à l'autre, d'un lieu à l'autre voire d'un auteur à l'autre. Ainsi, en faisant un test avec l'intégralité des productions indexées dans le WoS en 2007 (moyenne mobile sur 3 ans), on trouve un ratio de 1,9 adresses par article en Italie et au Japon pour 1,3 adresses par article en Russie et en Nouvelle-Zélande. Si la France et les États-Unis sont parmi les pays dont le ratio est le plus élevé, celui-ci ne s'élève qu'à 1,8. Contrairement à ce qui est communément avancé, cette valeur suggère que la France n'est pas le pays où la fragmentation institutionnelle est la plus forte. Pour s'en assurer, il faudrait toutefois vérifier que le résultat se confirme lorsqu'on ne considère que les articles écrits par un seul auteur. En effet, puisque ces chiffres ne sont pas pondérés par le nombre d'auteurs, rien n'indique que nous ne devrions pas interpréter ce résultat par la plus ou moins grande propension des chercheurs à collaborer suivant les pays et les disciplines (en sociologie par exemple, il a été montré que les auteurs américains ont généralement plus de co-auteurs que les auteurs français, voir Pontille, 2003).

Finalement, puisque selon le mode de comptage utilisé, le nombre d'adresses par article peut être défavorable ou favorable à l'impact des articles, cette question est associée de près à un enjeu stratégique, celui de la normalisation des pratiques de signatures pour la mesure et l'évaluation de l'activité scientifique d'une part ; celui de la simplification du paysage institutionnel local en matière d'enseignement supérieur et de recherche

d'autre part. En dehors de la nature stratégique de l'enjeu associée à la mise en compétition des lieux de science, il est clair que s'il existait une norme et si l'organisation administrative des établissements scientifiques était harmonisée, les résultats des études scientométriques seraient plus simples à interpréter.

En effet, l'exemple des Pays-Bas montre qu'il y a bien des façons de regrouper l'information et de la compter. Les choix qui sont faits en matière de regroupement et de comptage influencent notablement les résultats, surtout au milieu de la hiérarchie. Pour autant, il n'y a pas de « bonne » manière de procéder car chaque méthode donne à voir un aspect différent de la production scientifique. Sachant cela, la situation actuelle pose problème à plusieurs égards :

1) Il y a souvent une grande opacité dans les travaux existants en scientométrie spatiale sur la méthodologie utilisée (Gauffriau *et al.*, 2008 ; Bornmann & Plume, 2011).

2) Il s'agit de subtilités au sujet desquelles les commentateurs des classements internationaux sont assez insensibles (Lin, Huang, & Chen, 2013). Le classement de Leyde constitue une exception car il est le seul qui offre la possibilité, lors d'une consultation en ligne, de sélectionner le mode de comptage (fractionné ou non) de son choix pour obtenir le score des universités.

3) Le remplissage des adresses par les auteurs n'est pas normalisé et les pratiques varient selon les auteurs, les disciplines, et probablement les lieux géographiques. La fragmentation institutionnelle au sein des lieux scientifiques et géographiques ajoute de la confusion pour les chercheurs qui se trouvent confrontés à plusieurs options pour déclarer leur affiliation et en particulier, à la possibilité de déclarer une ou plusieurs adresses.

1.3. Les politiques de normalisation des adresses

Regardant ce troisième point, la parution du classement de Shanghai à partir de 2003 ou *Academic Ranking of World Universities* (ARWU) a mis en évidence le fait que pour certains systèmes de recherche, comme le système français, l'indexation des contributions universitaires dans les bases de données bibliographiques souffrait d'une absence de normes. Prenant conscience de la confusion que cela pouvait créer et en particulier, des problèmes de visibilité de la recherche associés à cette confusion, des chartes de normalisation à l'usage des chercheurs ont commencé à circuler en France pour la rédaction des adresses. L'Observatoire des Sciences et Techniques (OST) à Paris a lancé le projet « NormAdresses » sur le libellé des adresses françaises (OST, 2007 ; Aubert, 2008, p. 16-18). Ce travail montre que la déclaration de leur affiliation par les chercheurs n'est pas le seul paramètre qui entre en jeu dans le référencement puisque l'affiliation est ensuite reprise par un éditeur scientifique, puis par un opérateur d'indexation comme l'ISI Thom-

son Reuters dans le cas du WoS (OST, 2007 ; Dassa *et al.*, 2014). C'est ensuite l'affiliation telle qu'elle est indexée dans la base de données qui est reprise par les scientomètres pour la fabrication d'indicateurs ou de classements. Ces filtres successifs ne garantissent pas à l'adresse déclarée par le chercheur de rester intacte au bout de la chaîne. Cela dit, si jamais l'adresse déclarée est claire, pas trop longue, et qu'elle suit l'ordre logique prévu par l'ISI Thomson Reuters (« organisation principale, sous-organisation, ville, province, pays »), il n'y a pas de raisons qu'elle soit modifiée surtout si l'université est déclarée comme « organisation principale ». Aussi, il est souhaitable que la graphie utilisée pour faire référence à l'organisation de recherche soit unifiée car la co-existence de plusieurs désignations est source d'erreur et de confusion au moment de traiter l'information.

Ayant pris conscience de cet enjeu, les établissements de recherche se chargent désormais d'imposer une graphie unique afin d'améliorer leur visibilité. Le problème auquel restent confrontés les chercheurs, en particulier les chercheurs français, au moment de déclarer leur affiliation, vient de leur appartenance à des laboratoires mixtes ou UMR, autrement dit à des laboratoires disposant de plusieurs tutelles (université et organismes de recherche). Le choix de déclarer une tutelle plutôt que l'autre ou bien une tutelle avant l'autre a jusqu'alors été de la responsabilité du chercheur. Le problème se complique avec la question du nombre de lignes à utiliser pour signer (c'est-à-dire du nombre d'adresses). Certaines institutions comme l'Institut National de la Santé et de la Recherche Médicale (INSERM) en France dans le cadre de l'Alliance nationale pour les sciences de la vie et de la santé (Aviesan) craignent d'être reléguées au second plan si jamais toutes les tutelles devaient être déclarées sur une seule ligne, les unes à la suite des autres⁶⁴. Cette crainte, que ne partage pas le Centre national de la recherche scientifique (CNRS), s'explique par la tendance actuelle à mettre en avant la tutelle universitaire et par le fait que les outils d'analyse bibliométrique automatisés comme l'*Essential Science Indicators* (ESI) commercialisés par l'ISI Thomson Reuters ne retiennent pour chaque adresse que le nom de l'organisation supposée « principale » dans l'adresse (OST, *op. cit.*).

Pour que les autres organisations soient visibles dans l'ESI ou d'autres classements fonctionnant sur le même principe, les chercheurs relevant de l'Aviesan sont donc incités à

⁶⁴ Aviesan. (2011). Charte des publications (Adresses - Affiliations) (3 p.). Aviesan. En ligne sur le site : <http://www.aviesan.fr>

Voir aussi : Haefner-Cavaillon, N. (2012, février). La bibliométrie comme outil d'évaluation et de pilotage à l'INSERM. Présenté à Groupe de réflexion sur l'analyse et la caractérisation de la production scientifique-Outils et Méthodes, Urfist de Paris.

Notes consultables à l'adresse <http://urfist.enc.sorbonne.fr/veille-et-recherche/ateliers/evaluation-caracterisation-bibliometrie/seance-du-12-fevrier-2012>.

déclarer une adresse par tutelle au lieu de préciser toutes leurs tutelles sur une seule ligne. Pour l'OST, l'inconvénient majeur de cette pratique, baptisée « multi-ligne », signalé dans le rapport de la Commission Aubert, est de faire grimper artificiellement les chiffres de collaborations scientifiques. Avec cette pratique, il devient impossible de distinguer une situation de collaboration scientifique d'une situation de multi-tutelle. Cela contribue à l'augmentation du nombre d'adresses par article, qui est une tendance lourde constatée depuis les années 1990 mais généralement interprétée comme le signe de l'augmentation des pratiques de collaborations scientifiques dans la longue durée.

Le monoligne et les effets sur l'indexation des traitements automatiques de l'ISI Thomson Reuters

Modèle mono-ligne

Extrait d'un article où les auteurs ont appliqué le modèle mono-ligne

T. Cadoudal¹, J. M. Blouin¹, M. Collinet¹, F. Fouque¹, G. D. Tan², E. Loizon³, E. G. Beale⁴, K. N. Frayn², F. Karpe², H. Vidal³, C. Benelli¹ and C. Forest¹ ✉

(1) INSERM UMR-S 530/747; Université Paris Descartes, Centre Universitaire des Saints-Pères, 45 rue des Saints-Pères, F-75006 Paris, France

(2) Oxford Centre for Diabetes Endocrinology and Metabolism, Churchill Hospital, University of Oxford, Oxford, UK

(3) INSERM, U449; INRA 1235; Université Claude Bernard, Faculté de Médecine R. Laennec, Lyon, France

(4) Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, Lubbock, TX, USA

Enregistrement correspondant dans le WOS

Thomson a réordonné les adresses au profit des universités. L'université Claude Bernard a été retranscrite sous l'appellation « Univ Lyon 1 ».

Addresses: Forest C (reprint author), Univ Paris Descartes, INSERM, UMR S 530 747, Ctr Univ St Peres, 45 Rue St Peres, F-75006 Paris, France
Univ Paris Descartes, INSERM, UMR S 530 747, Ctr Univ St Peres, F-75006 Paris, France
Univ Oxford, Churchill Hosp, Oxford Ctr Diabet Endocrinol & Metab, Oxford, England
Univ Lyon 1, Fac Med R Laennec, **INSERM**, U449, **INRA** 1235, F-69365 Lyon, France
Texas Tech Univ, Hlth Sci Ctr, Dept Biochem & Cell Biol, Lubbock, TX 79409 USA

Lancement de la fonction « Analyze » sur cette publication





VIEW RECORDS	Field: Institution Name	Record Count	% of 1	Bar Chart	SAVE ANALYSIS DATA TO FILE
<input type="checkbox"/>	TEXAS TECH UNIV	1	100.0000 %		
<input type="checkbox"/>	UNIV LYON 1	1	100.0000 %		
<input type="checkbox"/>	UNIV OXFORD	1	100.0000 %		
<input type="checkbox"/>	UNIV PARIS DESCARTES	1	100.0000 %		
VIEW RECORDS					SAVE ANALYSIS DATA TO FILE

Figure 34 – Le monoligne et les effets sur l'indexation des traitements automatiques de l'ISI Thomson Reuters. Extrait du rapport d'étape du groupe de travail « Analyse des Instructions » pour le projet Norm-Adresse (OST, 2007)

Le problème ne s'arrête pas là pour l'OST car la pratique du « multi-ligne » a également des conséquences sur l'impact des publications en cas de comptage fractionné, qui est le mode de comptage retenu par l'OST pour produire les indicateurs bibliométriques requis, en France, par la mise en œuvre de la LOLF (la Loi organique relative aux lois de finance). En cas de fractionnement et d'autant plus si le comptage est de type « complet », le score de la publication est partagé par toutes les adresses. Selon l'intérêt qui est en jeu, cette division du score est plus ou moins heureuse. Ainsi, s'il s'agit uniquement d'améliorer le score des universités françaises dans l'ESI, il vaut mieux n'utiliser qu'une seule adresse en déclarant l'institution universitaire en premier. S'il s'agit au contraire de mettre en avant le poids d'un organisme comme l'INSERM, il est préférable de le déclarer en premier ; mais comme l'illustre la Figure 34, si une université figure également dans l'adresse, il y a de grande chance pour que l'algorithme de l'ISI Thomson Reuters réordonne automatiquement l'adresse au profit de l'institution universitaire ; ce qui du coup, justifie l'addition d'une ligne supplémentaire ne signalant que l'INSERM.

Mais est-ce bien raisonnable de vouloir pousser les chercheurs à adapter leurs pratiques de signature à la logique des routines informatiques de l'ISI Thomson Reuters alors que techniquement, il est tout à fait possible de tenir compte de toutes les institutions signalées dans une adresse « mono-ligne » ? Comme le résume l'OST en prenant l'exemple d'un article signé avec le modèle mono-ligne : « Dans un produit comme ESI ou dans la fonction "Analyze" du WoS, cette publication ne sera jamais attribuée à l'INSERM ou à l'INRA bien que ces organismes soient dans les tutelles des UMR car ils sont identifiés comme des "*sub-organizations*". Par contre, pour les indicateurs de la LOLF, le repérage fait dans le cadre de l'OST à partir de la base complète de Thomson Scientific permet d'attribuer ces publications aux bons opérateurs grâce à un travail minutieux de repérage des organismes et des universités dans toutes les adresses. » (OST, *ibid.*).

Indépendamment de la question de la place (dont il ne faudrait pas oublier de discuter dans les arènes où l'évaluation est faite) d'un institut privé comme l'ISI Thomson Reuters dans l'administration du monde scientifique, l'idéal reste de toute façon que les pratiques de signature soient harmonisées mais aussi les pratiques des scientomètres pour faciliter la comparabilité et au-delà l'interprétation que l'on peut faire des données bibliométriques. En ce sens, les initiatives de ces dernières années ne sont pas tout à fait convaincantes puisque les organismes, en France en tout cas, n'ont pas réussi à se mettre d'accord sur une façon unique de procéder : le CNRS est d'accord pour apparaître après l'université sur une seule ligne, alors que l'INSERM souhaite que les auteurs lui consacrent une ligne indépendante de la ligne où figure l'université qui les héberge. À cette mésentente, s'ajoute celle qui provient de la constitution de nouveaux ensembles comme les Pôles de Recherche et d'Enseignement Supérieurs (PRES) devenus Communautés

d'universités et d'établissements (COMUE). Ces derniers rassemblent sous une même appellation « Université » tous les établissements d'enseignement supérieur et de recherche d'une aire géographique donnée. Désormais, les chercheurs sont incités à déclarer leur appartenance à ces structures en plus de leur appartenance institutionnelle classique. Cette nouvelle pratique permettrait de faire gonfler artificiellement le score des « Universités » françaises dans les classements universitaires (notamment parce que l'ISI Thomson Reuters associe cette appellation à l'idée d'« organisation principale »). Nos observations suggèrent que la complexité de ce problème vient de la multiplicité des intérêts en présence et il semble que l'intérêt national soit actuellement de mettre en avant le poids des universités car ce sont les organismes scientifiques les plus reconnus à l'international. Dans un contexte où au lieu de classer les universités, ce seraient les organismes de recherche qui seraient mis en compétition, alors d'autres recommandations s'appliqueraient.

L'intérêt de la méthode de comptage entier que nous avons privilégiée vient de ce qu'elle permet de rester relativement peu sensible à ce qui se passe à l'intérieur du périmètre local. En effet, les entités spatiales qui participent à une publication reçoivent la même part de contribution qu'il y ait une seule adresse ou 50 adresses qui soient comprises dans le périmètre des entités spatiales participantes. Aussi, travailler au niveau de l'agglomération permet de simplifier l'information et de s'affranchir de la complexité interne au milieu local. Le choix de cette méthode est en plus justifié par notre approche conceptuelle : on sait que l'intensité relationnelle est forte à l'intérieur des agglomérations mais on préfère laisser cette réalité de côté pour s'intéresser à l'organisation mondiale, à la situation des agglomérations dans un contexte national et international. C'est là tout l'intérêt du principe de généralisation de l'information que décrit Denise Pumain :

« un objet formé de constituants hétérogènes à un niveau d'agrégation géographique peut paraître homogène à un niveau supérieur, parce qu'on oublie volontairement l'hétérogénéité interne pour ne retenir que les traits qui en permettent la comparaison avec d'autres unités spatiales situées au même niveau d'observation. » (Pumain & Saint-Julien, 1997, p. 37).

2. Le traitement des données de collaboration

« Chaque élément peut le plus souvent, être pris à son tour comme un système dont l'environnement est formé par les relations entre les autres éléments. » (Brunet & Dollfus, 1990, p. 282)

2.1. Du réseau de co-signatures au réseau de lieux

Dans les réseaux de collaboration scientifique généralement analysés en bibliométrie, l'unité d'analyse est le chercheur et l'on s'intéresse au nombre de publications produites conjointement, c'est-à-dire co-signées, par les chercheurs. En l'occurrence, puisque c'est l'organisation mondiale qui nous intéresse, le but est de capturer les tendances lourdes qui pèsent sur les collectifs de recherche au niveau mondial. Compte tenu du volume de données bibliographiques qu'il faut traiter dans cette perspective, le réseau de chercheurs qu'il serait possible d'en déduire est extrêmement volumineux. Un réseau de cette taille gagne à être simplifié pour être analysé. De plus, étant donné la hiérarchie entre auteurs telle qu'elle est reflétée par les publications (par exemple, en biomédecine, le chef d'équipe signe tous les articles de l'équipe en dernière position), travailler sur le réseau mondial des collaborations au niveau des auteurs compliquerait inutilement l'objet de notre étude en mettant en évidence les structures hiérarchiques internes aux équipes de recherche (organisations en étoile). Selon notre approche géographique, les publications des auteurs appartenant à une même entité géographique sont agrégées. Une fois cette opération d'agrégation de l'information réalisée, il est possible de construire des matrices relationnelles de lieux par année et d'en extraire des graphes de relations entre entités géographiques par année.

En mathématiques, un graphe est un ensemble de points (appelés nœuds ou sommets) qui peuvent être reliés entre eux par des liens. Le recours à ce type d'objets mathématiques est utile pour explorer et analyser des jeux de données relationnelles. Il est justifié par une grande variété de questions de recherche et on l'associe le plus souvent à des interrogations portant sur la notion de « réseau ». En général, les tenants de l'analyse de réseaux appellent « réseaux » les graphes que l'on peut extraire des données du monde réel. Cette notion, qualifiée par certains de « fourre-tout », est mobilisée par ceux qui mettent en avant le rôle des structures pour expliquer les phénomènes (autrement dit, postulant que la structure d'un phénomène renseigne son essence). En géographie, cette notion est associée plus spontanément à l'étude des réseaux de villes, réseaux d'infrastructures ou réseaux migratoires ; en sociologie et économie : réseaux d'individus

ou réseaux d'organisations ; en sociologie historique : réseaux d'évènements ; en biomédecine : réseaux de neurones ou réseaux de gènes ; en épidémiologie : réseaux de contacts ; en linguistique et science de la communication : réseaux de mots ou de documents. Les objets de prédilection des chercheurs varient d'une discipline à une autre mais il serait réducteur de considérer que chaque chercheur doit se limiter à interroger une catégorie spécifique de réseaux relevant plus naturellement de son champ de compétence. La littérature en témoigne, il arrive que des réseaux de mots soient analysés par des géographes (Beauguitte, Guérin-Pace, & Richard, 2014), des réseaux migratoires par des historiens (Lemerrier & Rosental, 2009), des réseaux de références par des sociologues (Milard, 2010) etc. Comme l'expliquait déjà Brian Berry en 1964 au sujet du rapport entre la géographie et les sciences sociales, ce n'est pas tant l'objet qui compte mais le point de vue (Berry, 1964).

L'analyse de réseaux permet de rendre compte des relations entre les éléments de la population à laquelle on s'intéresse. S'en servir suppose de faire un effort de formalisation consistant à définir rigoureusement quels sont les éléments de cette population (les nœuds ou sommets) et quelle est la nature de la relation supposée structurante qui les unit (les liens, arêtes ou arcs). Les liens traduisent les relations d'appartenance des sommets à des sous-ensembles inclus dans l'ensemble formé par la population du graphe. Dans un graphe de collaborations scientifiques par exemple, si les nœuds sont des chercheurs, ils appartiennent tous à l'ensemble que constitue la « communauté scientifique » et sont, en ce sens, tous liés. Un graphe classique de collaboration rend compte de l'appartenance des chercheurs à des sous-ensembles définis par le fait de partager une propriété supplémentaire en commun : avoir co-écrit un article scientifique. Dans l'approche que nous avons mise au point pour étudier la géographie des collaborations scientifiques, les nœuds sont des entités géographiques et les liens entre deux entités se voient attribuer une valeur selon le nombre de publications co-signées depuis ces entités. Cette approche permet de situer directement l'analyse dans le registre de la géographie : on identifie les co-publications de l'échantillon par leur lieu de production et non par leurs auteurs. On analyse un réseau de lieux et non un réseau d'individus.

Il existe plusieurs options pour attribuer une valeur aux liens de co-signatures. Tout comme nous fractionnons les publications pour mesurer des stocks de publications par entité spatiale, il est possible de fractionner la valeur des liens de collaboration pour générer des matrices de collaboration entre entités spatiales. Pour chaque co-publication, la valeur associée aux liens entre entités spatiales dépend alors du nombre d'entités spatiales ayant participé à la publication. Cette méthode de comptage est une transposition aux collaborations du « comptage entier fractionné » que nous utilisons pour mesurer des stocks de publications par agglomération (*Whole Normalized Counting*). En l'occurrence,

elle permet de donner plus d'impact aux collaborations deux à deux. Le postulat de cette démarche est que l'importance scientifique de la relation est proportionnelle au nombre de partenaires qui ont collaboré pour produire la publication. Si n est le nombre d'entités spatiales signataires pour une publication donnée, la valeur l du lien de collaboration unissant deux à deux les entités spatiales vaut :

$$l = \frac{1}{\frac{n(n-1)}{2}}$$

Ainsi, la somme de la valeur des liens de collaboration est égale au nombre total de publications écrites en collaboration entre au moins deux entités spatiales. Pour reprendre l'exemple de la publication produite conjointement par 16 chercheurs répartis entre les agglomérations de Toulouse et Paris (Figure 31), la relation entre Toulouse et Paris sera créditée de 1 pour cette publication. Si une agglomération supplémentaire avait été partie prenante de cette production, le lien entre Toulouse et Paris aurait été crédité de $1/3$; pour 4 agglomérations impliquées, ce lien vaudrait $1/6$ etc. Comme pour le traitement des données de production, le choix d'un comptage entier au détriment d'un comptage complet revient à traiter les entités élémentaires d'analyse, en l'occurrence les agglomérations, comme des boîtes noires. Lorsqu'on mesure la participation des agglomérations avec cette méthode, il n'est pas question de se soucier du nombre d'adresses par agglomération, et *a fortiori* du nombre d'auteurs, qui ont participé à chaque publication. Pour chaque publication, le seul critère important (unité de compte) est le nombre d'agglomération qui ont participé, peu importe le nombre d'adresses, d'auteurs ou d'organisations par agglomération qui se sont impliqués dans la publication. Cette méthode de comptage est adaptée à notre parti pris méthodologique puisque l'appliquer revient à s'intéresser aux liens entre les lieux sans se soucier de la dynamique relationnelle interne aux lieux, autrement dit, à l'écologie des lieux.

À plus d'un titre, les matrices d'adjacence obtenues en appliquant cette méthode sont une simplification de l'information relationnelle qui est contenue dans les données bibliographiques. Compte tenu de la structure de la base de données, les lieux ne sont reliés entre eux que parce qu'ils sont associés à des publications communes. Or, comme l'explique Antoine Grandclément dans sa thèse, il existe une étape intermédiaire avant d'arriver à la matrice de lieux, c'est l'étape où nous disposons d'une table relationnelle entre notices bibliographiques et lieux de science (Grandclément, 2012, p. 200-207). D'une telle table, il est possible d'extraire un graphe bipartite ou *two-mode* (publications \times lieux) pouvant ensuite être converti en graphe *one-mode* (lieux \times lieux). Cette conversion que tous les logiciels d'analyse de réseaux permettent de réaliser est le résultat d'une opé-

ration matricielle classique (le produit de la transposée de la matrice de publications par elle-même). Pour beaucoup de spécialistes en analyse de réseaux, il importe d'être conscient de la perte d'information induite par cette opération (Ducruet, 2010). À cet égard, le schéma ci-dessous, extrait de la thèse de Grandclément, tient lieu de démonstration (Figure 35).

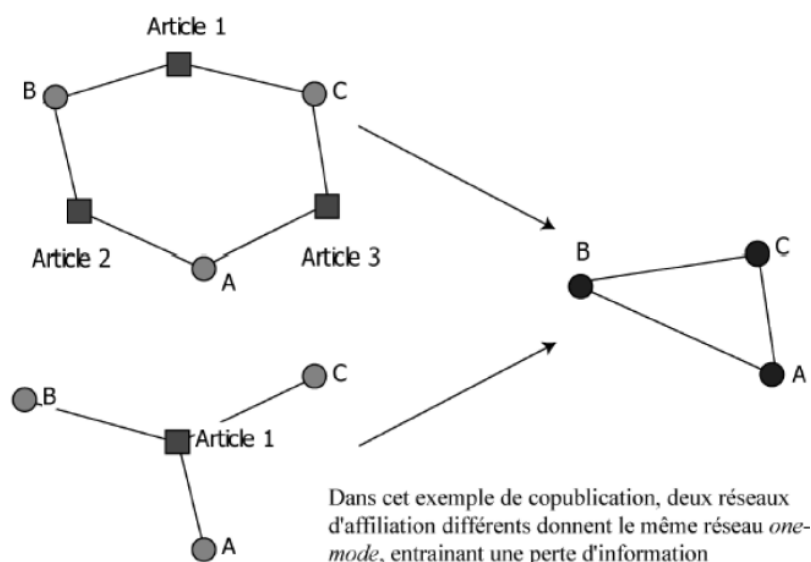
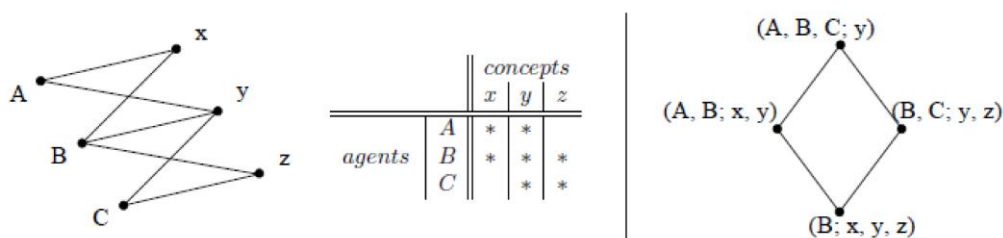


Figure 35 – La perte d'information dans la transformation d'un réseau *two-mode* en réseau *one-mode*. Schéma extrait de la thèse d'Antoine Grandclément (Grandclément, 2012, p. 200-207)

Il se trouve que les graphes bipartites se prêtent bien aux données bibliographiques parce qu'ils permettent de considérer plusieurs populations à la fois et qu'en bibliométrie, étant donné la richesse de l'information contenue dans une notice bibliographique, il peut être intéressant de considérer les relations entre entités appartenant à différentes populations. Parmi les populations envisageables se trouvent celles des publications, des mots-clefs, des auteurs, des lieux, ou encore des références. Dans sa thèse de sociologie sur les « communautés épistémiques », Camille Roth a ainsi travaillé à des méthodes d'analyse de graphes bipartites associant des auteurs à des concepts scientifiques. Il a même envisagé l'opportunité d'avoir recours à un type de formalisation rarement mobilisé en sciences sociales : le treillis de Galois (Roth, 2005, p. 23-31). La Figure 36 montre le processus permettant de passer d'un graphe bipartite à un treillis de Galois. Dans son étude, le treillis permet d'ordonner des communautés épistémiques (CEs) en allant des plus générales (davantage d'agents, moins de concepts) aux plus spécifiques (moins d'agents, plus de concepts) (Roth, 2008).



À gauche : réseau socio-sémantique montrant les liens entre agents (A, B et C) et concepts (x, y et z). Au milieu : matrice équivalente (l'astérisque indique la présence ou non de lien entre un agent et un concept). À droite : treillis correspondant, regroupant les quatre CEs, de la plus générale (en haut) à la plus spécifique (en bas).

Figure 36 – Le passage d'un graphe bipartite à un treillis de Galois. Schéma extrait de l'article « Réseaux épistémiques : formaliser la cognition distribuée » (Roth, 2008)

En fait, les spécialistes en analyse de réseaux ou structures complexes, y compris les plus géographes d'entre eux, ont très tôt été sensibles à la perte d'information induite par le recours aux graphes simples (*one-mode*) pour formaliser des phénomènes observés dans le réel. Dès 1980, Peter Gould reprenant les travaux du mathématicien Ron Atkin, regrettait que la théorie des graphes soit trop pauvre pour rendre compte de la multi-dimensionnalité du monde et au-delà de la complexité des structures : « Mais les structures sont si complexes dans les systèmes naturels et humains que même le graphique le plus beau et sophistiqué ne suffit pas à en rendre compte. »⁶⁵ (Gould, 1980). Pour remédier à ce problème, il propose avec Jeffrey Johnson de recourir à la *Q-analysis* ou *Polyhedral Dynamics* développée par Ron Atkin dans les années 1970 (Gould & Johnson, 1980). Cette formalisation permet de raisonner avec au moins trois dimensions. Son apport théorique majeur consiste en la distinction entre le décor (*backcloth*) et le trafic (*traffic*). Dans une pièce de théâtre, le décor contraint le déroulement de l'action sur la scène (s'il y a un obstacle, les acteurs doivent le contourner). Pour autant, le décor n'est pas fixe : il peut être modifié pendant le spectacle et alors modifier le cours de l'action tandis que l'action elle-même est en mesure de le transformer. Appliqué à l'exemple classique du réseau de co-signatures entre auteurs, l'appartenance à la « communauté scientifique » fait partie du décor et le fait de collaborer représente le trafic. L'ensemble obtenu n'est pas stable dans le temps : il est possible d'intégrer ou de quitter la communauté au cours du temps tandis que les collaborations sont des événements ponctuels qui relèvent du trafic. *A priori*, tous

⁶⁵ L'original : « But structural matters can become so complex in natural and man-made systems that even sophisticated and beautifully drawn graphics begin to break down. The basic problem is the multidimensional nature of complexity, and the fact that the objects of our inquiries eventually force us to abandon the limited two dimensions of a page, and even the third dimension we can capture by playing a trick on our eyes with perspective. » (Gould, 1980).

les membres de la communauté sont susceptibles de collaborer entre eux et chaque publication co-produite par deux membres ou plus peut être traitée comme un fait structurant.

Avec la *Q-Analysis*, chacun de ces faits peut être formalisé sous la forme d'un polyèdre (un ensemble de un, deux, trois... n dimensions). Si pour chaque publication, la dimension du polyèdre dépend du nombre de signataires par publication, il est possible de concevoir un « complexe simplicial » qui donnerait à voir la structure obtenue en connectant les polyèdres partageant des auteurs en commun, autrement dit les publications communes à plusieurs membres. Il y a alors fort à parier qu'en procédant ainsi pour un corpus de publications en biomédecine, on mettrait en évidence la structuration des équipes collaborant régulièrement au sein de chaque laboratoire. Dans la logique d'Atkin, l'analyse de réseaux gagne en profondeur si elle est effectuée au regard d'une structure sous-jacente qu'il est possible de formaliser à l'aide d'un « complexe simplicial ». En 1980, tandis que Peter Gould publie une introduction à la *Q-analysis* pour les géographes et urbanistes (*ibid.*), les sociologues Patrick Doreian et Linton C. Freeman, intéressés par l'évolution des réseaux sociaux, l'expérimentent. Tous deux concluent qu'il n'y a rien qu'ils aient pu faire avec la *Q-analysis* qui n'aurait pu être exprimé par la théorie des graphes (Doreian, 1979 ; Freeman, 1980). En revanche, ils insistent sur le caractère heuristique de la distinction entre décor et trafic. Aujourd'hui, les analyses de graphes bipartites se sont substituées à la *Q-analysis* d'Atkin, laquelle n'est plus tellement utilisée en sciences sociales. En 1999, Narciso Pizarro introduit les notions de « places » et de « réseaux de places » (Pizarro, 1999). Le raisonnement est proche de celui d'Atkin : le réseau de places est obtenu à partir d'un graphe bipartite et s'apparente au décor d'Atkin. Ces approches s'adaptent bien à l'appréhension des réseaux en dynamique (Maisonobe, 2013a).

La difficulté à laquelle est soumis le chercheur face à ce type de formalisation vient du fait que les outils opérationnels dont il dispose pour l'analyse sont moins nombreux et que les résultats obtenus sont plus délicats à interpréter (Beauguitte, 2013, *op. cit.*). La différence de lisibilité entre la représentation d'un graphe *one-mode* et celle d'un graphe *two-mode* suffit à en témoigner. Le graphe *one-mode*, qui traite des relations existantes au sein d'entités appartenant toutes à la même population, est beaucoup plus simple à analyser et à visualiser. Il faut néanmoins garder à l'esprit qu'extraire un graphe *one-mode* d'une base de données relationnelle suppose de sacrifier un grand nombre de dimensions. Dès lors, en s'intéressant directement et uniquement à la matrice d'échange entre lieux sans passer par la table relationnelle associant les lieux aux publications, nous renonçons à la possibilité de saisir des combinaisons de lieux (d'une dimension supérieure à 2). Pour le dire autrement, cette réduction suppose d'appréhender les collaborations scientifiques

uniquement à travers des dyades, c'est-à-dire des liens deux à deux, visualisables sur un plan en deux dimensions.

En l'espèce, il importe de retenir qu'en formalisant l'information sous forme de matrices d'échange entre agglomérations scientifiques et en appliquant un comptage entier fractionné des liens, nous renonçons à analyser :

- les connexions entre agglomérations et publications même si, et c'est là un grand avantage de la méthode retenue, nous gardons une trace de ces relations à travers la valeur associée aux liens entre agglomérations puisque le fractionnement se fonde sur le nombre total d'agglomérations participant à chaque co-publication ;

- les connexions entre adresses appartenant à la même agglomération puisque le comptage entier traite à égalité toutes les agglomérations participant à une publication quel que soit le nombre d'adresses par agglomération. Il élimine ainsi la possibilité de considérer des liens réflexifs (boucles) sauf si l'on passe à un niveau supérieur à celui de l'unité spatiale de base retenue pour les calculs, l'agglomération.

Ainsi, seuls les liens entre adresses appartenant à des agglomérations différentes sont traités comme des collaborations scientifiques. Puisque les collaborations internes à une institution et à une agglomération ne sont pas comptabilisées, un grand nombre de co-signatures sont ignorées par cette méthode. Bien que ce choix soit justifié par la nécessité de simplifier l'information pour l'analyser, il est important d'y revenir pour en illustrer les conséquences et en discuter la pertinence d'un point de vue géographique. Une fois que les agglomérations de localités publiantes ont été constituées, la méthode privilégiée pour étudier les données de collaboration implique que les liens entre localités appartenant à la même agglomération soient ignorés au profit des liens entre agglomérations. En procédant de la sorte, il est indispensable d'avoir conscience du rôle que jouent les périmètres d'agrégation sur la structure des réseaux de collaboration obtenus à la sortie. Justement, la première étude de cas réalisée au cours de notre cursus doctoral a été l'occasion d'évaluer et de réfléchir à la portée de ce parti pris méthodologique.

2.2. Agrégation spatiale versus cohésion scientifique

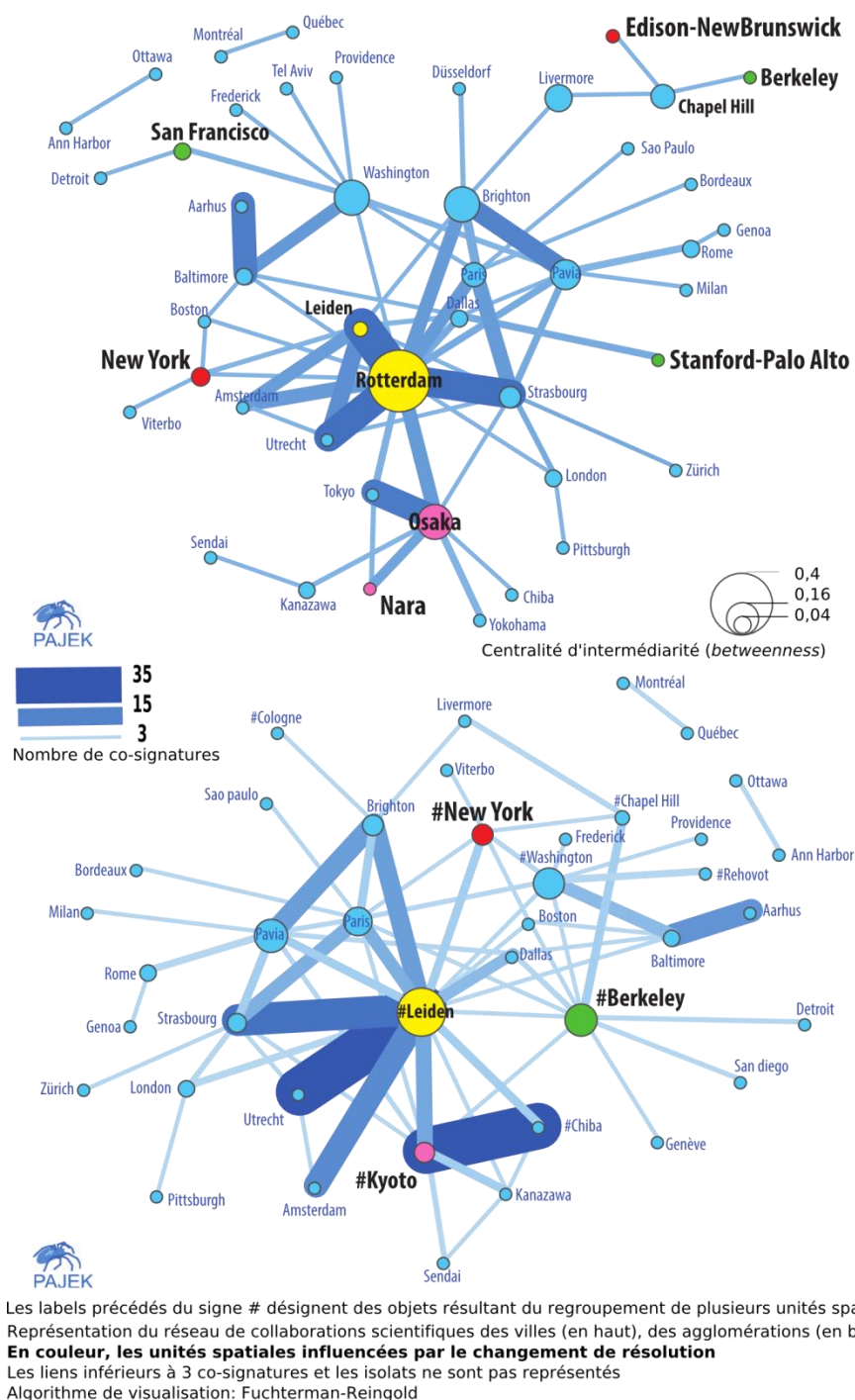
Les périmètres d'agglomération définis dans le Chapitre 6.2.2 ont été pensés pour s'adapter aux traitements géographiques et en dynamique de jeux de données bibliométriques au niveau mondial. Comme ils sont insensibles aux maillages administratifs et à leurs évolutions, ces périmètres d'agrégation sont supposés opérationnels pour comparer le niveau de production scientifique des lieux de science dans l'espace et le temps, au moins pour la période allant des années 1990 aux années 2010. Les mobiliser pour étudier

la géographie d'une spécialité scientifique a permis de les mettre à l'épreuve ainsi que la méthode de synthèse de l'information au fur et à mesure de leur élaboration.

Pour étudier la géohistoire d'un collectif de recherche en biomédecine, plusieurs corpus bibliographiques ont été analysés. Présentée dans la dernière partie de la thèse, cette recherche a permis de révéler plusieurs aspects de la géographie de la spécialité de la réparation de l'ADN. Le premier (dans l'ordre chronologique) corpus de publications analysé dans cette perspective porte sur un problème précis apparu dans les années 1990 sur le lien entre un mécanisme de réparation de l'ADN et un facteur de transcription de l'ADN. Une fois ce corpus extrait du *Web of Science* (WoS), il a fait l'objet d'une analyse spatiale (Partie 3.Chapitre 11.1). Pour cette analyse, plusieurs périmètres locaux d'agrégation des données ont été considérés de telle sorte que les périmètres d'agglomération utilisés pour obtenir les premiers résultats présentés au Colloque Européen de Géographie Théorique et Quantitative (ECTQG) en 2011 à Athènes (Baron, Eckert, & Maisonobe, 2011) ne sont pas les mêmes que les périmètres utilisés pour la version définitive de l'étude de cas publiée dans la revue *M@ppemonde* en 2013 (Maisonobe, 2013b). L'influence sur le réseau mondial des collaborations entre villes résultant de ce changement de périmètre est digne d'intérêt. Plus précisément, il importe de voir dans quelle mesure ce changement peut affecter la structure du réseau de collaboration scientifique obtenu à partir du corpus de publications analysé.

À cet effet, la Figure 37 a été réalisée qui montre deux représentations graphiques du même réseau de co-publications. Un jeu de données identique est utilisé mais les périmètres d'agrégation des données sont différents. Sur la représentation du haut, les entités sont les villes et leur proche banlieue (premier niveau d'agrégation : 45 nœuds). Ces périmètres sont ceux que nous avons conçus en 2011, avant que des périmètres définis à partir d'un protocole stabilisé aient été définis par l'équipe du programme « Géoscience » : « les agglomérations scientifiques ». En dessous, il s'agit du même réseau mais construit au niveau des agglomérations scientifiques fabriquées avec la méthode « Géoscience » (second niveau d'agrégation : 40 nœuds). Sur cette représentation graphique, la taille des cercles est proportionnelle à l'indice de « centralité d'intermédierité » qui permet de mettre en évidence la qualité d'entremetteur occupée par les nœuds dans le réseau (pour une définition plus précise, voir le Chapitre 8.2.2). En couleur, figurent les nœuds dont nous avons voulu souligner la sensibilité au changement de périmètre.

**Collaborations scientifiques autour de la question
du lien entre transcription et réparation de l'ADN**
L'effet sur la structure du réseau de copublications (1990-2007)
de l'agrégation des données à différents niveaux géographiques
L'exemple du passage des villes aux agglomérations scientifiques



Conception et réalisation: Marion Maisonneuve. Retouché par Laurent Jégou pour publication dans M@ppemonde.

Figure 37 – L'effet sur la structure du réseau de co-publications (1990-2007) de l'agrégation des données à différents niveaux géographiques. Source : Web of Science

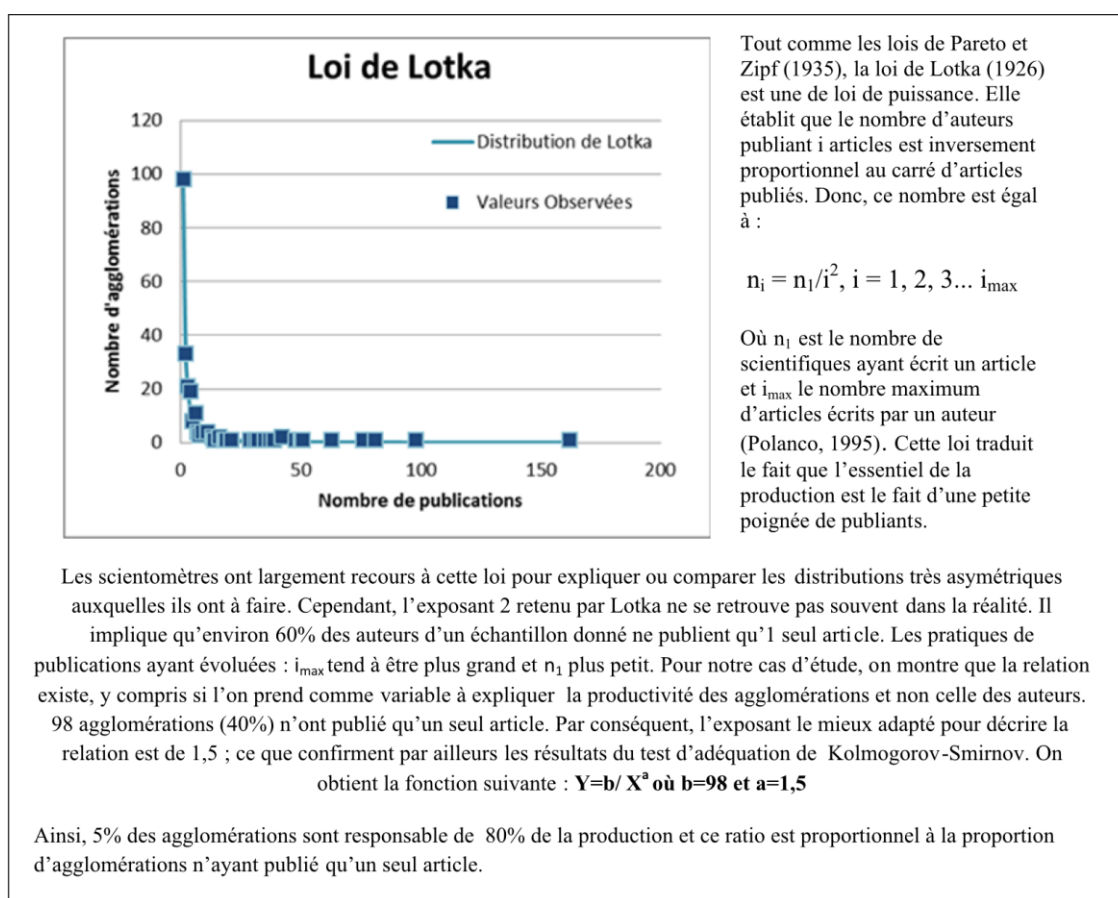
Seulement deux cas ont un effet remarquable sur la structure du réseau : le regroupement de Stanford, San Francisco et Berkeley (en vert) et de New York avec Edison et Brookhaven (en rouge). Ces changements de périmètres améliorent mécaniquement la connexion des villes américaines. Avec le découpage en « agglomération scientifique », l'agglomération de New York (#New York) et celle de la baie de San Francisco (#Berkeley) prennent une place centrale dans le réseau de ville ; alors même que la pertinence de ces périmètres n'est pas confirmée par la complicité scientifique unissant les équipes qui s'y trouvent. En revanche, l'entité qui regroupe les équipes de Leiden (Leyde en français) et Rotterdam aux Pays-Bas (en jaune) et celle qui regroupe les équipes d'Osaka et Nara au Japon (en violet) sont conformes à l'état de leur connivence scientifique regardant la question étudiée. Compte tenu de l'intensité des collaborations à l'intérieur de ces périmètres, les regroupements sont à la fois cohérents géographiquement et scientifiquement.

Les collaborations scientifiques étant un bon indicateur de la structuration de la question de recherche au niveau international, le fait que deux localités géographiquement proches ne co-publient pas est une information importante sur la cohésion scientifique locale regardant cette question. Cependant, la cohésion scientifique est une donnée très sensible, surtout au niveau d'une question de recherche. En tenir compte pour délimiter les périmètres d'étude pourrait conduire à remettre en cause certains regroupements intra-institutionnels ou intra-municipaux qui ne seraient pas forcément sanctionnés par une collaboration scientifique ou bien encore à faire évoluer les périmètres au gré du développement de nouveaux partenariats scientifiques.

Un exemple tiré de l'étude de cas illustre ce phénomène. Il concerne le lien entre deux organismes de recherche américains en matière de réparation de l'ADN. D'après nos analyses, le *National Cancer Institute* (NCI) de Bethesda travaille beaucoup avec le *National Institute of Aging* (NIA) sur le lien entre transcription et réparation de l'ADN. Les deux institutions sont rattachées administrativement au *National Institute of Health* (NIH). Au début des années 1990, elles sont toutes deux situées à Bethesda (dans l'État du Maryland, en banlieue de Washington). Or, au milieu des années 1990, le NIA déménage à Baltimore ce qui se traduit par la mise en place de collaborations régulières entre Baltimore et Bethesda. Un système de cours par vidéoconférence est lancé à cette occasion qui ne tarde pas à être rendu accessible aux autres centres du NIH sur le territoire américain, puis mis en ligne pour un accès généralisé (Kraemer & Bohr, 2005). Ce faisant, le déménagement a contribué à accentuer l'intensité des échanges scientifiques entre les deux villes voisines. En conservant le même périmètre d'agrégation des données dans le temps, nous pouvons même évaluer l'importance de cette évolution. Mais ce déménagement intervenu au cours de la période d'étude suffit-il à justifier le regroupement de Baltimore et Washington dans une même entité d'analyse ?

Si le NIA avait déménagé à l'autre bout des États-Unis, l'intensité de la collaboration n'aurait pas forcément été différente puisque les deux centres bénéficiaient à l'avance d'une forte complicité scientifique et de moyens technologiques permettant de faciliter leur collaboration. En fait, si la problématique était celle de la structuration scientifique de la question de recherche (sans considération pour sa géographie), il serait pertinent de regrouper ensemble tous les centres du NIH dans une même entité (quelle que soit leur localisation). Cela permettrait de situer le NIH par rapport aux autres organismes scientifiques travaillant sur ce sujet de recherche. À cette fin, on regrouperait aussi ensemble tous les centres britanniques du *Medical Research Center* (MRC) et toutes les unités de l'Institut Français de Recherche Médicale (l'INSERM). Dans ce cas, l'entité d'analyse ne serait plus l'entité spatiale mais l'organisme de recherche. Comme on l'a déjà vu dans le Chapitre 6, ce n'est pas que cette approche soit sans intérêt mais elle ne convient pas pour une étude prenant la localisation géographique des activités scientifiques comme variable pour analyser leur dynamique. En l'espèce, ce qui compte pour étudier l'interaction des lieux d'activité scientifique au niveau mondial c'est davantage la stabilité des entités spatiales et leur homogénéité que leur cohésion scientifique regardant la question de recherche ou même la spécialité étudiée. Cette dernière est beaucoup moins stable dans le temps puisqu'elle peut être affectée par des déménagements ou d'autres événements imprévisibles comme le départ à la retraite d'un chef d'équipe.

Ainsi avons-nous tiré de cette étude de cas la conviction que lorsque c'est le processus de diffusion spatiale de la question de recherche qui est placé au centre de l'analyse, le fait que les entités considérées soient homogènes dans le temps et l'espace est plus important que le fait que les chercheurs de toutes les localités qui s'y trouvent collaborent. C'est d'autant plus vrai dans le cas de l'analyse d'une question de recherche puisque, dans ce cas-là, parler des publications d'une agglomération revient généralement à parler des publications de l'unique équipe scientifique intéressée par la question dans cette agglomération. En effet, dans le corpus étudié, une agglomération est égale à une équipe/institution dans près de 70 % des cas. Par conséquent, la plupart du temps, le passage au niveau de l'agglomération ne change rien aux résultats obtenus. La comparaison des représentations graphiques commentées ci-dessus montre bien que seul un petit nombre de zones géographiques sont affectées par les choix qui sont faits en matière de découpage des agglomérations. Ces découpages délicats sont ceux qui concernent les quelques zones caractérisées par une forte densité de publications et d'affiliations. En effet, seulement quelques zones posent problème parce que les publications du corpus sont extrêmement concentrées. Le corpus obéit à la loi de Lotka dans la mesure où 80 % des publications sont produites par 5 % des agglomérations (Encadré 2).



Encadré 2 - La loi de Lotka. Encadré extrait de l'article « Diffusion et structuration spatiale d'une question de recherche en biologie moléculaire » (Maisonobe, 2013b)

Dans le même temps, seulement 70 agglomérations sur 229 (soit 30 %) cumulent plusieurs affiliations différentes (nombre d'adresses supérieur à un). Autrement dit, les zones à forte densité de publications ainsi que les zones à forte densité d'affiliations sont rares à l'échelle mondiale. Il reste que le rang (en nombre de publications) des lieux situés en zone dense, mais surtout leur place dans le réseau de co-publications, sont considérablement influencés par le découpage géographique retenu pour passer des localités aux agglomérations. Au-delà de l'étude de cas, c'est ce qui explique que ces zones sensibles aient pu servir à notre équipe pour les arbitrages à faire pour le découpage des « agglomérations scientifiques ».

Pour appliquer le critère de densité de population à l'échelle mondiale, il a effectivement fallu définir un seuil au-delà duquel tracer les limites des agglomérations situées en zones denses. Pour déterminer le seuil à partir duquel délimiter les zones à forte densité de population, la Mégalopole nord-américaine a servi de zone-test à notre équipe de recherche. Il s'agissait de déterminer un seuil de rupture dans la densité de population qui

permettrait de tracer des limites entre Washington, Baltimore, Philadelphie, Princeton, New York, New Haven, Providence et Boston. Le seuil retenu pour délimiter des taches urbaines dans cet espace urbain et scientifique majeur a ensuite été appliqué pour guider le découpage manuel des 376 « agglomérations scientifiques » construites à partir du critère démographique. C'est ainsi que les périmètres obtenus présentent l'avantage de ne pas tenir compte de la fragmentation administrative et de sa variabilité en fonction des pays. Comme ces périmètres ont été délimités à l'aide d'un protocole unique fondé sur un critère démographique pour les localités les plus productives et leurs proches voisines et un seuil de distance kilométrique pour les autres, ils permettent la comparaison internationale (Chapitre 6).

Pour autant, il n'est absolument pas question de sacraliser ces périmètres. Puisqu'à chaque étude de cas doit correspondre une question différente, des besoins spécifiques peuvent très bien justifier des réaménagements ponctuels tout à fait conciliables avec l'approche de « scientométrie spatiale » défendue ici. Sachant qu'il n'existe pas de périmètres parfaitement homogènes, il importe surtout d'avoir conscience de l'effet des découpages sur les résultats des analyses. Les « agglomérations scientifiques » permettent de simplifier l'information mais incitent à laisser de côté des informations potentiellement intéressantes sur les liens entre proches localités. C'est ainsi que même en travaillant à l'échelle mondiale et à condition que la recherche le justifie, il est parfois pertinent de regarder en détail des cas situés en « zones sensibles ». C'est ce que nous avons fait lors de l'étude de la question « réparation et transcription de l'ADN » en interrogeant la nature des relations effectives entre Washington et Baltimore. Au-delà, bien que nous nous soyons posée la question d'adapter les découpages à la « cohésion scientifique » (c'est-à-dire à la densité des échanges scientifiques) intra-agglomération, c'est finalement les périmètres des « agglomérations scientifiques » construites en équipe qui ont été adoptés pour l'étude de cas. En l'espèce, l'idée de tenir compte des relations formelles (les liens de co-signature) entre les chercheurs spécialisés dans la question très précise du lien entre réparation et transcription de l'ADN pour délimiter des périmètres d'analyse a très vite été évacuée par la rareté des liens effectifs entre localités publiantes sur ce thème.

En prenant uniquement le problème par ce biais, nous n'avons pas fait le tour de la question. En effet, élaborer un critère de cohésion scientifique pour délimiter des entités spatiales offre plusieurs possibilités. Au lieu de s'en tenir à la cohésion scientifique dans le domaine de la réparation de l'ADN, il serait sans doute plus pertinent de définir une cohésion scientifique toutes disciplines confondues et sur le temps long. À part éventuellement le caractère non-exhaustif des données disponibles sur les interactions scientifiques, rien n'indique que ce critère serait moins efficace que celui de la fréquence des échanges domicile-travail pour délimiter des espaces fonctionnels ou « territoires ».

Il nous semble que plusieurs points de vue sont tenables sur l'intérêt et la faisabilité d'un tel projet et qu'il mériterait de plus amples investigations. Les points de vue exprimés par Marie-Noëlle Comin et Christian Matthiessen *et al.* à cet égard montrent qu'un consensus n'a pas émergé quant au type de périmètres à mobiliser pour étudier les activités scientifiques.

Matthiessen explique que négliger la cohésion sociale interne aux entités spatiales est le prix à payer pour travailler avec des périmètres homogènes d'un point de vue géographique : « Cette façon de délimiter des unités d'une manière plutôt spatiale, en appliquant une règle stricte pour établir les unités, nous conduit à combiner des villes équivalentes avec d'importantes interactions (par exemple, les unités de l'aire Rhin-Rhur), mais à mettre aussi ensemble des villes très différentes avec de faibles interactions (par exemple, Oxford et Reading). Nous avons été attentifs à ne pas trop laisser notre différent niveau de connaissance sur les villes influencer cet exercice. »⁶⁶ (Matthiessen, Schwarz, & Find, 2002). À l'inverse, Comin accepte l'idée de faire quelques « petites entorses à la règle du seuil de distance imposé [qui] sont d'une part, d'un ordre de grandeur minime (qui ne dépasse pas la dizaine de kilomètres en moyenne) ; et d'autre part ont été effectuées pour un nombre très restreint de localités (82), pour lesquelles la pertinence de l'agrégation à l'aire urbaine fonctionnelle de l'agglomération a été systématiquement contrôlée et ne faisant aucun doute. » (Comin, 2009, p. 213-214). Elle justifie ainsi le rattachement de Fiskebäckskill (à 70 km de Göteborg) et Lysekil (à 10 minutes en ferry de Fiskebäckskill) en Suède à l'aire urbaine de Göteborg par leurs collaborations effectives avec l'université de Göteborg. De même, elle justifie le rattachement de Font-Romeu-Odeillo-Via (à 72 km) à l'aire fonctionnelle de Perpignan par le lien institutionnel entre le four solaire d'Odeillo et la filière STAPS de cette commune avec l'université de Perpignan. Ces choix sont légitimés par le fait qu'elle a choisi, comme entité élémentaire d'analyse, l'aire fonctionnelle ou d'influence des grands centres scientifiques.

Cette dernière approche serait délicate à mettre en application à l'échelle mondiale car elle suppose de se documenter sur un très grand nombre de cas particuliers à moins, comme nous le suggérons implicitement plus haut, de concevoir une méthode d'agrégation automatisée mobilisant les données de collaborations scientifiques effectives entre les localités géocodées. La difficulté majeure de ce projet est qu'il exige d'accorder une confiance importante à la qualité du géocodage. À vrai dire, ce n'est pas que l'idée de tenir compte des relations scientifiques pour construire des périmètres d'étude ne mérite

⁶⁶ L'original : « This way of delimiting units in a rather spatial manner, and using a strict rule to aggregate the units thus established, combines similar cities with extensive interactions (for example, the units of the Rhine-Ruhr area), but it also puts together dissimilar cities with few interactions (for example, Oxford and Reading). We have been careful not to let our different level of knowledge of the cities influence the exercise. » (Matthiessen, Schwarz, & Find, 2002)

pas d'être creusée ; c'est juste que pour des analyses exploratoires et préliminaires à l'échelle mondiale, la solution consistant à agréger les localités géocodées au sein d'entités construites à l'aide d'un critère démographique au niveau desquelles généraliser l'information bibliographique s'est révélée préférable et plus évidente à mettre en œuvre.

En somme, bien que la méthode générique mise au point dans le cadre de « Géoscience » se soit révélée opérante, elle ne prétend pas être l'unique méthode de traitement de données bibliographiques qui s'impose pour une géographie des sciences. Il s'agit d'une méthode adaptée à l'analyse de la répartition géographique annuelle d'ensembles de publications scientifiques ainsi qu'à l'analyse des réseaux de collaborations qu'il est possible d'en déduire à l'échelle mondiale au niveau de résolution géographique le plus fin possible. Au-delà, et en particulier pour nos recherches sur la géographie d'une spécialité scientifique, cette méthode générique de traitement a été reprise mais aussi complétée, nous le verrons dans la dernière partie de la thèse, par d'autres approches et d'autres types de traitements bibliométriques.

Avant d'en venir à la présentation des résultats généraux et dépasser l'étape de la méthodologie, il convient de s'y attarder dans un dernier chapitre portant sur les méthodes de l'analyse des données proprement dite. Nous commençons par le rôle de la visualisation dans le travail d'analyse et terminons par une présentation des principaux outils d'analyse quantitative retenus pour étudier la géographie mondiale de la production scientifique contemporaine.

CHAPITRE 8. L'ANALYSE SPATIALE DES DONNÉES BIBLIOGRAPHIQUES

Après avoir défini un niveau élémentaire d'analyse et sélectionné un mode de comptage des publications et des collaborations, nous pouvons disposer des variables suivantes pour tout ensemble de notices bibliographiques :

- Une liste de lieux à partir de l'adresse des auteurs ;
- Un nombre de publications scientifiques par lieux, et par année ;
- Un nombre de relations entre ces lieux compte tenu des co-publications, par année.

C'est seulement à ce stade que nous pouvons interroger les variables, les explorer en vue de fournir des réponses pour, enfin, en savoir plus sur la géographie mondiale des activités scientifiques. Il faut rappeler que les choix méthodologiques et les périmètres d'agrégation des données qui ont été définis jusqu'ici sont calibrés pour des analyses de la production scientifique à l'échelle mondiale et plus précisément pour des analyses spatiales horizontales c'est-à-dire pour la comparaison des lieux et de leur position dans l'espace et le temps en matière de science. Les moyens d'exploiter ou de mettre à profit de telles observations sur l'activité scientifique sont nombreux. Encore une fois, si nous ne présentons ici que les moyens les plus appropriés pour un travail de type exploratoire sur la géographie des activités scientifiques à l'échelle mondiale, il ne fait aucun doute que d'autres outils et techniques d'analyse peuvent se justifier.

Premièrement, nous considérons l'opportunité de visualiser les données et montrons qu'obtenir une représentation graphique lisible et interprétable demande un effort supplémentaire de simplification de l'information qui relève de l'analyse. Dans ce cas, la représentation graphique possède les vertus d'un résultat de recherche puisqu'elle est le produit fini auquel on parvient une fois que les données à analyser ont traversé une série de filtres successifs permettant de les rendre intelligibles et leur donner une portée scientifique. À cet égard, une attention particulière est accordée aux solutions d'exploration visuelle et interactive des données et leur utilité à différents stades du travail de recherche.

Deuxièmement, nous montrons l'intérêt d'avoir recours aux outils de la statistique descriptive et de l'analyse de réseaux pour explorer et analyser la géographie des productions scientifiques et des collaborations scientifiques à l'échelle du monde. Les statistiques permettent de raisonner en relatif et de mesurer des évolutions et donc de repérer des

tendances concernant la répartition des activités scientifiques et leur organisation réticulaire. L'analyse de réseaux permet d'interroger la structure géographique qui se dégage des réseaux de lieux scientifiques, et de caractériser la place de chacun des lieux dans un système mondial.

1. La visualisation et la représentation des données :

« Mais le spectacle n'est pas identifiable au simple regard, même combiné à l'écoute. Il est ce qui échappe à l'activité des hommes, à la reconsidération et à la correction de leur œuvre. Il est le contraire du dialogue. Partout où il y a représentation indépendante, le spectacle se reconstitue. » (Debord, 1967, p. 23)

1.1. La visualisation de « big data » à l'échelle mondiale

La visualisation et le traitement de l'information scientifique ont connu des développements spectaculaires au cours des vingt dernières années. Les figures tracées à la main en 1994 par John F. Padgett et republiées il y a quelques années dans la revue REDES rappellent combien est récent le recours systématique à l'informatique pour le traitement de données quantitatives (Padgett, 2011). Aujourd'hui, nous disposons de logiciels spécialisés dans l'exploration et l'analyse des données (parmi d'autres, on peut citer les logiciels libres abordés plus loin : R, Quantum GIS, Inkscape), qui, manipulés avec une certaine expérience, produisent des résultats de grande qualité, tant sur le plan scientifique que sur celui de l'aspect visuel. Cette nouvelle dimension de la recherche mérite d'être approfondie pour que le passage à la visualisation reste un outil précieux de l'analyse et de l'argumentation. Il existe désormais, pour répondre à cette nécessité, des experts en visualisation d'informations. Un exemple : Jean-Daniel Fekete (INRIA) travaille avec des chercheurs de différentes disciplines pour affiner les usages possibles des logiciels et effectuer les meilleurs choix en matière de visualisation des données⁶⁷. Ce processus est incontournable quand le volume de données à analyser dépasse l'entendement (on parle alors de « big data »). L'image offre dans ce cas un moyen de synthétiser l'information en la simplifiant. Représenter l'information sur un plan en deux dimensions oblige forcément à sacrifier certaines dimensions. Mais récemment, l'engouement des

⁶⁷ Cf. : Beauguitte, L., & Fekete, J.-D. (2012, octobre). La parole à Jean-Daniel Fekete [Carnet de recherche du groupe fmr]. Consulté à l'adresse <http://groupefmr.hypotheses.org/626>

physiciens, informaticiens et autres spécialistes des systèmes complexes pour les « *big data* » a contribué à améliorer la prise en compte simultanée du temps et de l'espace. Ainsi, à titre d'exemple, la modélisation spatio-temporelle est devenue l'objet d'un nombre croissant de thèses (52 thèses ont pour mot-clef « modélisation spatio-temporelle », en France en 2005-2007, contre 336 en 2012-2014 : résultat d'une requête réalisée sur le moteur de recherche en ligne Thèses.fr le 02/07/2015). En dépit de ces évolutions, l'une des références en matière de sémiologie graphique, y compris pour les informaticiens comme Jean-Daniel Fekete, reste Jacques Bertin, cartographe de formation. C'est même sans doute grâce à lui que les géographes sont les chercheurs en sciences sociales parmi les plus sensibilisés aux questions de visualisation. Ils sont particulièrement concernés par la géovisualisation qui, grâce à de nouveaux moyens informatiques, évolue au sein d'un domaine de recherche à part entière, qui propose des méthodes de représentation efficaces, interactives et innovantes (MacEachren & Kraak, 2001). Mais l'intérêt pour la visualisation touche aussi d'autres sciences sociales. En particulier, avec le tournant quantitatif en sociologie et en histoire, et le développement des logiciels d'analyse de réseaux, la visualisation des données assistée par ordinateur a progressé dans ces deux disciplines (Orford, Harris, & Dorling, 1999 ; Lemercier, 2005 ; Lemercier, Guzzi-Heeb, & Bertrand, 2011).

L'analyse de réseaux et son pendant mathématique, la théorie des graphes, permettent, nous l'avons déjà évoqué, de scruter des données relationnelles. Un graphe admet plusieurs types de représentations graphiques parmi lesquelles la représentation matricielle et le diagramme nœud-lien (Henry & Fekete, 2008 ; Bahoken, 2011 ; Bahoken, Beauguitte, & Lhomme, 2013a). Ce dernier correspond à la représentation topologique du graphe et il est souvent désigné directement par le terme « graphe ». Il s'agit du mode de visualisation privilégié des analystes de réseaux sociaux. Les lignes symbolisent les relations. Elles sont plus ou moins épaisses, orientées ou non au moyen d'une flèche, et relient entre eux des points qui symbolisent les unités de base de l'analyse : cas, éléments ou individus. Malgré le fait que les analystes des réseaux sociaux se concentrent sur des réseaux d'individus, rien n'oblige à ce que les points du diagramme ou « nœuds » soient des individus, et les lignes ou « liens », la matérialisation des rapports interpersonnels qu'ils partagent à un temps donné. Aussi, quand on choisit de se focaliser sur la dimension temporelle d'un phénomène, il est tout à fait permis de définir un graphe qui, bousculant les habitudes, décrirait des relations entre des événements dans le temps ou bien des changements d'état (par exemple, un réseau de trajectoires individuelles, voir Figure 77, p. 375).

Si l'on se donne pour objectif de développer des photographies du « monde scientifique » contemporain, autant la carte que le diagramme nœud-lien présentent un intérêt en tant que *médium* ou appareils. Ces deux dispositifs sont même particulièrement adaptés

pour visualiser l'information géographique extraite des notices bibliographiques. Puisque l'activité scientifique est localisable, il est tout naturel d'avoir recours à la cartographie pour en rendre compte. Parallèlement, puisque nous disposons d'informations relationnelles à travers les co-signatures de publication, un mode de visualisation non spatialisé des relations peut également se justifier. En effet, pour analyser la structure ou l'organisation d'un « monde », la visualisation cartographique, qui tient compte de la distance kilométrique entre lieux dans l'espace géographique, n'est pas la seule digne d'intérêt. Représenter des liens sur une carte revient à donner beaucoup d'importance à la distance physique entre les lieux qui sont en relation. Visuellement, les liens entre lieux très éloignés sont davantage mis en valeur, y compris lorsqu'on fait varier l'épaisseur des liens pour exprimer la valeur associée aux liens ou, pour le dire autrement, l'intensité de la relation. Or, si c'est justement l'intensité relationnelle entre différents lieux qu'il s'agit de visualiser, il n'est pas forcément nécessaire de tenir compte de la distance physique qui les sépare. L'intensité de la relation entre les lieux peut très bien servir de métrique pour le placement des lieux, permettant de faire ressortir la structure sociale en question.

Le diagramme nœud-lien est intéressant pour obtenir l'image simplifiée d'une telle structure. L'intérêt de ce diagramme en géographie et sa complémentarité avec un mode de visualisation cartographique des informations est d'ailleurs en passe de devenir une évidence. Plus particulièrement, au cours des 10 dernières années, le problème de la représentation graphique des jeux de données relationnelles à l'échelle mondiale a pris de l'importance en géographie. Il faut dire qu'en dehors de l'évolution interne à la discipline, l'accès en ligne à des ensembles de données inédits sur les flux mondiaux associé aux facilités de mises en carte par ordinateur ont permis la circulation sur la toile d'images spectaculaires qui ne sont pas toujours du goût des plus fervents défenseurs des conventions cartographiques.

L'exemple archétypal est celui de la carte des amitiés Facebook publiée à la fin de l'année 2010 sur l'espace internet de la compagnie⁶⁸ (Figure 38). Réalisée à l'aide du logiciel R, cette carte a été abondamment commentée par les géographes dont Thierry Joliveau qui y a consacré plusieurs chroniques sur son site « Monde géonumérique » au début de l'année 2011⁶⁹. Pour Joliveau, il ne s'agit pas de nier la prouesse technique ni la qualité

⁶⁸ Butler, P. (2010, décembre 13). Visualizing Friendships. Consulté à l'adresse <https://www.facebook.com>

⁶⁹ Joliveau, T. (2011, janvier 10). 500 millions d'amis, la carte de Facebook – 1) Déconstruction. Consulté à l'adresse <https://mondegeonumerique.wordpress.com>

Joliveau, T. (2011, février 1). 500 millions d'amis, la carte de Facebook – 2) Analyse globale. Consulté à l'adresse <https://mondegeonumerique.wordpress.com>

Joliveau, T. (2011, avril 24). 500 millions d'amis, la carte de Facebook – 3) Évolution. Consulté à l'adresse <https://mondegeonumerique.wordpress.com>

esthétique de cette carte mais de la renvoyer à ce qu'elle est vraiment à savoir un objet idéologique : « C'est une construction faite pour communiquer efficacement la doctrine Facebook que l'on décrypte en creux : la compagnie relie les hommes du monde entier par-delà les frontières, les montagnes et les océans. Elle prend soin de toutes les amitiés qu'elle aide à éclore. Elle accompagne chacune d'entre elles avec attention. Elle connaît chaque fil particulier du grand écheveau rassurant et bienveillant qu'elle tisse autour de la planète grâce à son savoir technique parfaitement maîtrisé. ».



Figure 38 – Les « amitiés » Facebook par Paul Butler

Parmi tous les choix sémiologiques discutables dont résulte cette carte, le simple fait que nous ne connaissions pas le seuil de discrétisation utilisé pour représenter les liens d'amitié entre lieux lui enlève tout crédit scientifique. Quand très peu de temps après la sortie de cette carte, une production cartographique du même acabit est signée par Olivier Beauchesne⁷⁰ (Figure 39), un employé de la société Science-Metrix, sur les collaborations scientifiques de 2005 à 2009, elle ne tarde pas à être mise en cause : « En effet, la reprise, quasi à l'identique, des choix de conception de la carte "Facebook", conduit à penser que le lecteur de la carte de la science mondiale est exposé à des risques similaires (et majeurs) de mésinterprétation du document comme du phénomène représenté (...) L'effacement d'éléments structurants (comme les États) sur la carte, la mise en avant des liens sans l'accompagner d'un commentaire explicatif digne de ce nom ne permet en rien d'aller vers une compréhension du phénomène. On est dans le domaine de la performance visuelle et esthétique non-dénuée d'arrière-plans idéologiques, ce qui peut nourrir notre

⁷⁰ Beauchesne, O. H. (2011, janvier 23). Map of scientific collaboration between researchers. Consulté 17 février 2015, à l'adresse <http://olihb.com/2011/01/23/map-of-scientific-collaboration-between-researchers/>

malaise. Car on peut déduire de cette carte "sans frontières" que les collaborations scientifiques sont le produit de la simple volonté d'individus (ou de laboratoires) de travailler ensemble, comme se structurent librement (en apparence) les amitiés individuelles sur Facebook. » (Baron, Eckert, & Jégou, 2011). L'efficacité cartographique du résultat laisse les géographes d'autant plus perplexes que : « ce complexe tableau de fils devrait permettre de visualiser les relations scientifiques entre villes, mais en pratique on y repère surtout les zones de forte densité de production de la science. » (*ibid.*). Depuis, très loin d'avoir tenu compte de ces critiques, le même Olivier Beauchesne, satisfait de l'impact médiatique rencontré par son image, a récidivé en proposant une version actualisée de la même représentation⁷¹.

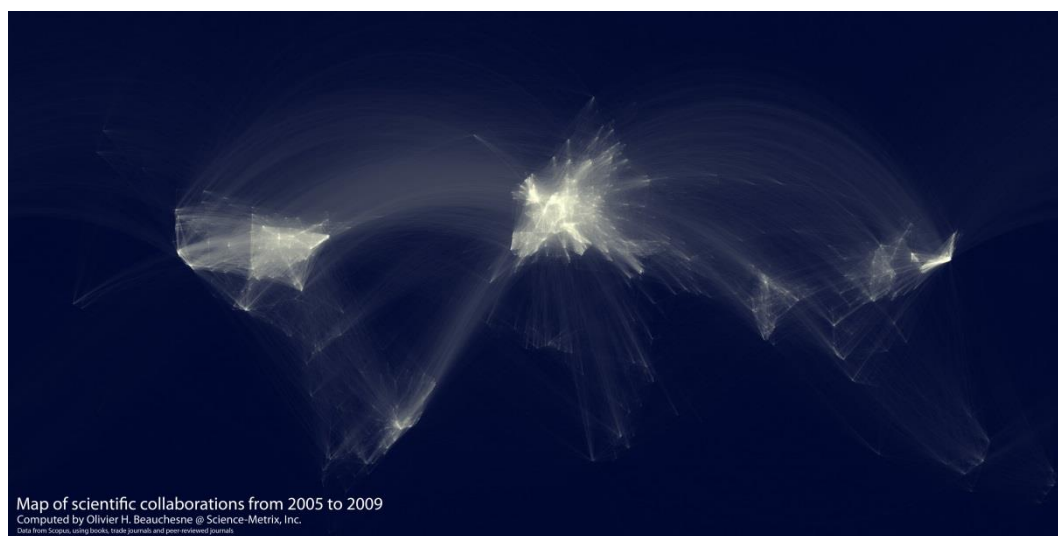


Figure 39 – Les collaborations scientifiques dans le monde par Olivier Beauchesne, 2011

Familier de ce qu'ils appellent les « cartes en oursins », les géographes ont tendance à les manipuler avec beaucoup plus de prudence. En effet, explique Joliveau : « Elles sont difficiles à lire dès que les lignes sont nombreuses et se croisent dans tous les sens (...) C'est encore plus compliqué quand on veut associer une quantité aux lignes. » (Joliveau, *op. cit.*). Produire une carte en oursin ou, pour utiliser le terme scientifique, une « carte de flux » efficace à l'échelle mondiale est donc un véritable défi. À l'occasion de l'édition 2013 de l'ECTQG à Dourdan, Bahoken *et al.* ont recensé et testé les divers procédés cartographiques permettant de réduire « l'effet spaghetti » sur ce type de carte : filtrage des liens, agrégations de l'information ou encore fusion des liens (*edge bundling*) (Bahoken, Beauguitte, & Lhomme, 2013b). Bien que ces procédés améliorent la lisibilité des cartes, il est préférable d'en avoir une utilisation rigoureuse et transparente puisque

⁷¹ Beauchesne, O. H. (2014, août 11). Map of scientific collaboration (Redux!). Consulté 17 février 2015, à l'adresse <http://olihb.com/2011/01/23/map-of-scientific-collaboration-between-researchers/>

comme ils le déduisent de cette expérience sur les cartes de flux mondiaux : « Si produire de jolies images est facile aujourd'hui, le véritable défi reste de produire du contenu scientifique pertinent. »⁷² (*ibid.*). Le respect des règles de sémiologie graphique et le recours à une légende, un titre et si possible un commentaire accompagnant la production sont essentiels dans cette perspective.

Comme évoqué plus haut, les géographes sont de plus en plus nombreux à explorer les opportunités offertes par d'autres modes de représentation des données relationnelles qui pourraient compléter utilement la traditionnelle carte de flux. Cette évolution tient au fait que plusieurs thèses de géographie ont eu à traiter d'ensembles de données dont pouvaient être extraits des réseaux de dimension mondiale au cours des dix dernières années. On compte parmi elles : celle de César Ducruet sur le trafic portuaire en 2004, celle de Maude Sainteville sur les flux financiers en 2009, et celle de Laurent Beauguette sur le vote des résolutions à l'ONU en 2011 (Ducruet, 2004 ; Sainteville, 2009 ; Beauguette, 2011). Ce dernier a d'ailleurs profité de son expérience doctorale pour lancer un groupe de réflexion sur la théorie des graphes avec César Ducruet : le groupe *fmr* (flux, matrice, réseau) auquel nous avons participé avec Françoise Bahoken, Serge Lhomme et Matthieu Drevelle. Parmi les membres du groupe, Bahoken est spécialisée dans la représentation des flux de population. Ses travaux montrent, tout comme les thèses précédemment évoquées l'illustrent, que les différents modes de représentation existants sont plus complémentaires qu'antagonistes. Pour Bahoken : « La carte de flux est une catégorie de diagramme de flux caractérisée par l'application d'une contrainte de force topologique sur le placement des lieux. » (Bahoken, 2011). Ainsi, si certaines contraintes de représentations sont communes à la carte et aux graphes (diagramme nœud-lien), d'autres exigent des ajustements spécifiques notamment, parce que « là où le graphe des flux privilégie l'affichage des liens sur les nœuds, la carte des flux privilégie celle des nœuds sur les liens » (*ibid.*). Dès lors, pour visualiser un réseau de lieux, utiliser un mode de représentation non spatialisé permet d'appréhender les lieux en situation (leur position relative et son évolution) sous l'angle de l'intensité relationnelle qu'ils entretiennent plutôt qu'en fonction de leur plus ou moins grand éloignement géographique. On se focalise sur la distribution des relations plutôt que sur la répartition spatiale des lieux.

En scientométrie, il existe aussi un sous-domaine de recherche en lien avec les questions de visualisation qualifié de *science mapping*. Parmi les protagonistes de ce champ de recherche, nous en avons déjà cité certains comme Felix de Moya Anegón (voir Partie 1. Chapitre 4.2.). Sans doute en raison de leur proximité cognitive avec les sociologues et contrairement aux géographes, ces spécialistes sont familiers des représentations de

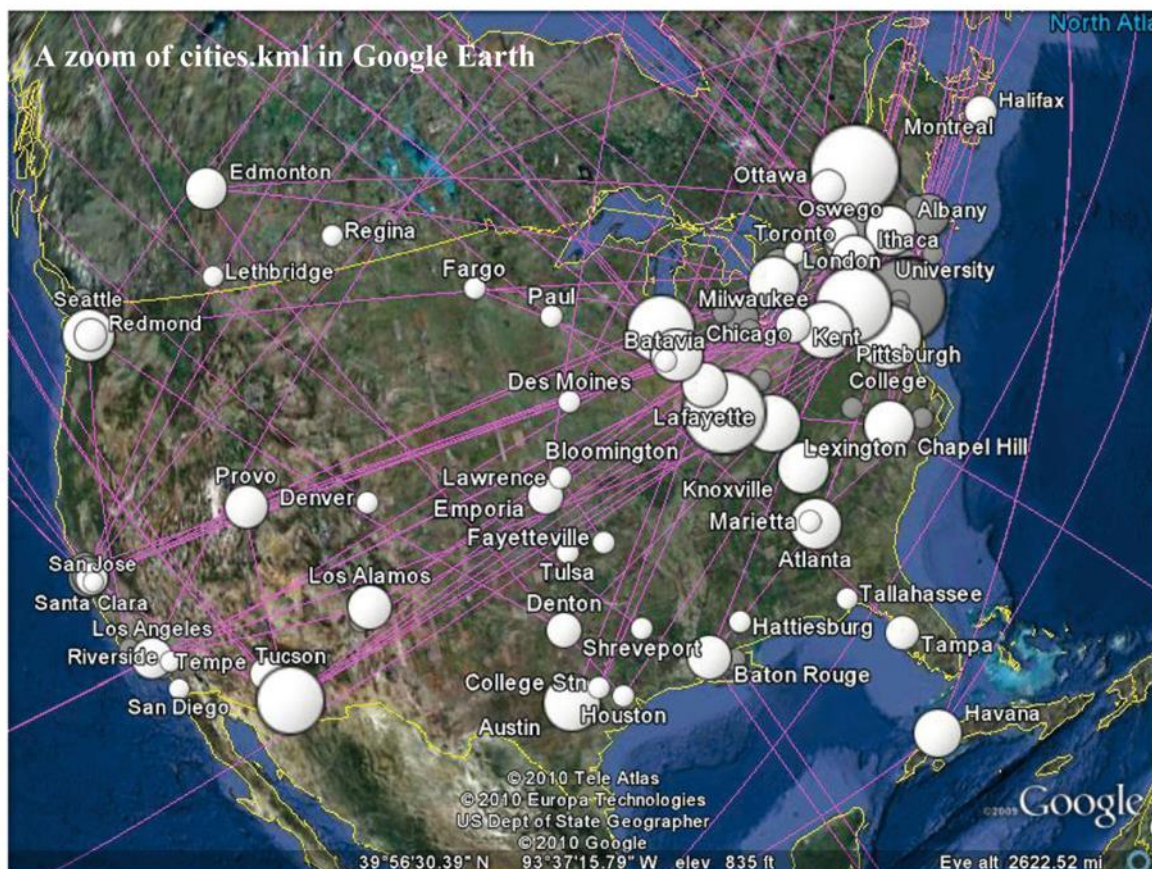
⁷² L'original : « If producing nice world pictures is easy today, the major issue remains to produce relevant scientific content. » (Bahoken *et al.*, 2013).

graphes non-planaires et non-spatialisés. Les réseaux qu'ils conçoivent sont de différentes natures (du réseau de co-signatures au réseau de co-citations en passant par le réseau de mots-clefs ou *coword*), mais quand il s'agit de traiter l'information géographique et, ainsi, de s'inscrire dans le domaine de la scientométrie spatiale, c'est plus souvent des réseaux de pays ou d'organisations qui sont représentés. La carte est apparue récemment en tant que mode de représentation dans ce domaine, mais les tentatives comme celles d'Olivier Beauchesne, soutenue d'ailleurs par Felix de Moya Anegón (Beauchesne, 2014, *op. cit.*), apparaissent symptomatiques de : « l'appropriation de l'outil cartographique par des acteurs qui n'ont pas fait l'effort de s'y former méthodologiquement et théoriquement. » (Baron *et al.*, *op. cit.*). Aussi, si l'on ne peut qu'applaudir le « caractère pionnier du travail de Loet Leydesdorff » dans l'élaboration de modules de géovisualisation des données bibliographiques sur *Google Earth* (Leydesdorff & Persson, 2010), il importe que les visualisations obtenues par ce biais soient mobilisées en tant que documents de travail ou comme outils pour l'exploration, mais certainement pas comme produits finis. En commentant une image publiée par Leydesdorff sur les collaborations scientifiques américaines (Figure 40), Eckert *et al.* justifient cette opinion : « les limites des utilitaires de géovisualisation apparaissent quand il s'agit d'aller vers une représentation cartographique de qualité. Ici *Google Earth* impose à l'utilisateur un fond graphiquement surchargé et aux informations inutiles (relief, bathymétrie...), avec des couleurs criardes aux limites de la lisibilité. La manière dont le logiciel gère l'affichage des noms complique encore la donne. Enfin, *Google Earth* ne crée pas de légende et ne permet pas de hiérarchiser le dessin des symboles pour éviter les superpositions. » (Eckert *et al.*, 2013). Il faut ajouter à l'appui de ces remarques, l'inconvénient de la projection de référence de *Google Maps* : la projection Mercator qui est un « objet anachronique et inadapté » à la représentation du monde (Jégou & Eckert, 2008).

Au bout du compte, de l'avis des géographes qui se sont penchés sur la question, trouver des solutions pour visualiser des données de flux à l'échelle mondiale est essentiel pour la discipline mais aussi pour les chercheurs en sciences sociales qui sont nombreux à se tourner vers une approche « globale » des phénomènes (Caillé & Dufoix, 2013). Dans cette optique, différents modes de représentation méritent d'être considérés. Surtout, il faut être attentif à accompagner chaque représentation, y compris les figures de réseaux non-spatialisés, de ses clefs de lecture (légendes, respect des règles de sémiologie graphique). Le logiciel libre de traitement d'images vectorielles Inkscape peut, par exemple, être mobilisé comme complément pour reprendre les sorties graphiques des logiciels d'analyse de réseaux⁷³.

⁷³ Cf. : Maisonnobe, M. (2014, 28 août). Atelier Pajek - Les réseaux. Atelier méthodologique présenté à Université d'été Ferney-Voltaire, Ferney-Voltaire. Consulté à l'adresse : groupefmr.hypotheses.org/3434

Une géovisualisation artisanale des données de production scientifique



L. Leydesdorff (2010)

Commentaire (D. Eckert)

Ce document "cartographique" que l'on peut consulter sur le site de L. Leydesdorff et qui a fait l'objet d'une publication dans JASIST (2011) est un exemple de certaines productions réalisées à tâtons avec Google Earth ou d'autres utilitaires de géovisualisation. Il peut frapper déjà par l'inutile surcharge du fond qui le rend très peu lisible, ainsi que par l'absence de procédure d'agrégation spatiale dans le tableau de données initial qui complique encore la donne. Conséquence: l'empilage de cercles dans la région de San Francisco ne laisse voir que le nom de "Santa Clara" et fait évidemment sous-estimer l'importance de cette région urbaine. Et pourtant, L. Leydesdorff s'est limité dans cet exemple à un seul domaine du WoS, les "Library and Information Sciences" : le nombre de lieux à cartographier est donc assez limité, et la « carte » résultante est assez simple.

Figure 40 – Une géovisualisation récente des données de production scientifique, via Google Earth avec l'aimable autorisation de l'auteur (© L. Leydesdorff) (Eckert et al., 2013)

En somme, au-delà de l'enjeu esthétique associé à toutes représentations graphiques, il faut être rigoureux pour que le produit de la visualisation puisse aspirer au statut de « résultat de recherche ». Nous allons voir, à présent, ce que cela suppose pour la représentation des données de production et de collaboration scientifique entre lieux.

1.2. La représentation comme résultat de recherche

Tandis que la visualisation assistée par ordinateur est aujourd'hui très utilisée en sciences sociales, il importe de ne pas confondre les méthodes de visualisation qui servent à l'étape d'exploration des données – comme, dans le cadre du programme « Géoscience », l'utilitaire de *wepmapping* conçu pour assister la vérification du géocodage et la construction des agglomérations scientifiques – avec les méthodes et concepts à appliquer pour produire des représentations graphiques ayant valeur et prétention de résultat scientifique. Cette dernière étape, soit parce qu'on ne peut pas tout montrer, soit parce que l'information dont on dispose est trop disparate, demande de faire un grand nombre de choix qui peuvent être discutés et qui sont forcément insatisfaisants parce qu'on perd de l'information. Celle-ci doit être simplifiée, car c'est ce qui la rend lisible et permet d'y associer un message. Dans ce qui suit, nous décrivons les représentations statiques puis les représentations interactives obtenues à l'issue du traitement géographique de l'intégralité du contenu du *Web of Science* (tous types de documents confondus) pour les séries d'années les plus récentes à notre disposition.

Pour aborder ce contenu, la carte a été privilégiée pour rendre compte de la répartition spatiale de l'activité scientifique et le diagramme nœud-lien (non-spatialisé) pour décrypter l'organisation de l'activité scientifique. Suivant ce parti pris, les Figure 41 et Figure 42 ont été conçues, en collaboration avec Laurent Jégou, pour être mises en regard avec d'autres images réalisées à partir d'observations scientométriques à d'autres époques et d'autres échelles : les chimistes européens au XVIII^e siècle par René Sigrist et les compagnons de banquet de Plutarque par Antony Andurand. Ainsi, nos deux représentations ont l'ambition de saisir un reflet du « monde scientifique » contemporain, celui de la « science visible » extraite du *Web of Science* (Andurand *et al.*, à paraître). L'emprise géographique de ce monde peut s'appréhender avec la carte et sa logique relationnelle avec le diagramme.

La carte montre les 200 agglomérations urbaines qui ont publié le plus en 2007. De ces 200 agglomérations proviennent 2/3 des publications indexées dans le *Web of Science* (WoS) en 2007 (environ 1 million d'articles sur 1,5 millions d'articles par an en moyenne mobile sur 3 ans). Elles relèvent de 42 pays différents. La représentation carto-

graphique permet de les localiser et de rendre compte du nombre de publications qui en proviennent (Figure 41).

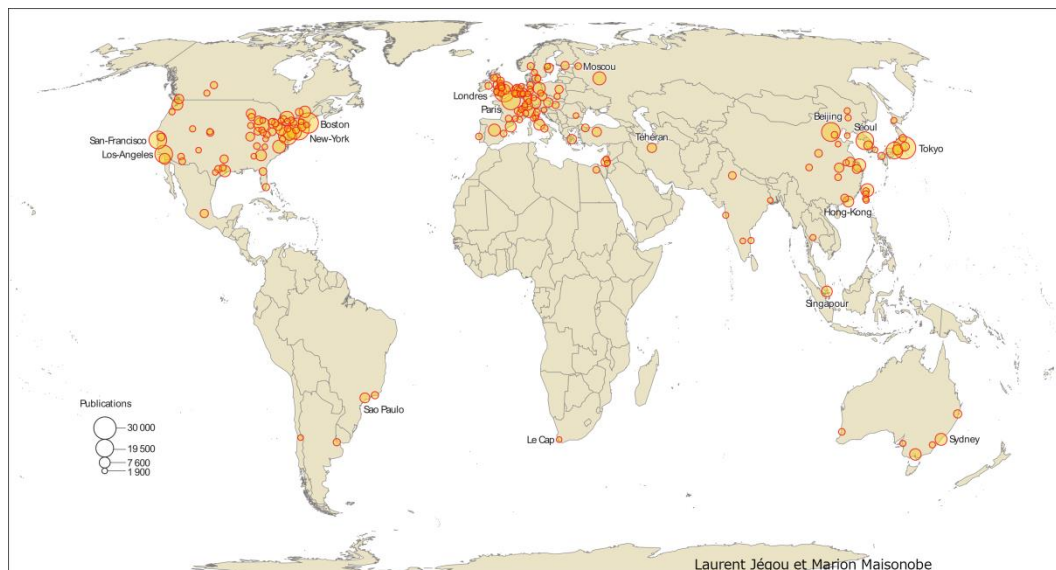


Figure 41 – Répartition géographique des publications scientifiques dans le monde en 2007.
Source : Web of Science, moyenne mobile sur 3 ans

Le résultat témoigne d'un faible niveau de publication en Afrique et dans une moindre mesure en Amérique du Sud. Bien que la carte n'en atteste pas, l'Asie est le continent dont l'activité « visible » a le plus progressé ces dernières années, si bien que, contrairement à ce qu'il se passait en 2000, la Chine compte, sur cette carte, autant d'agglomérations que le Royaume-Uni et l'Allemagne (soit 15 agglomérations).

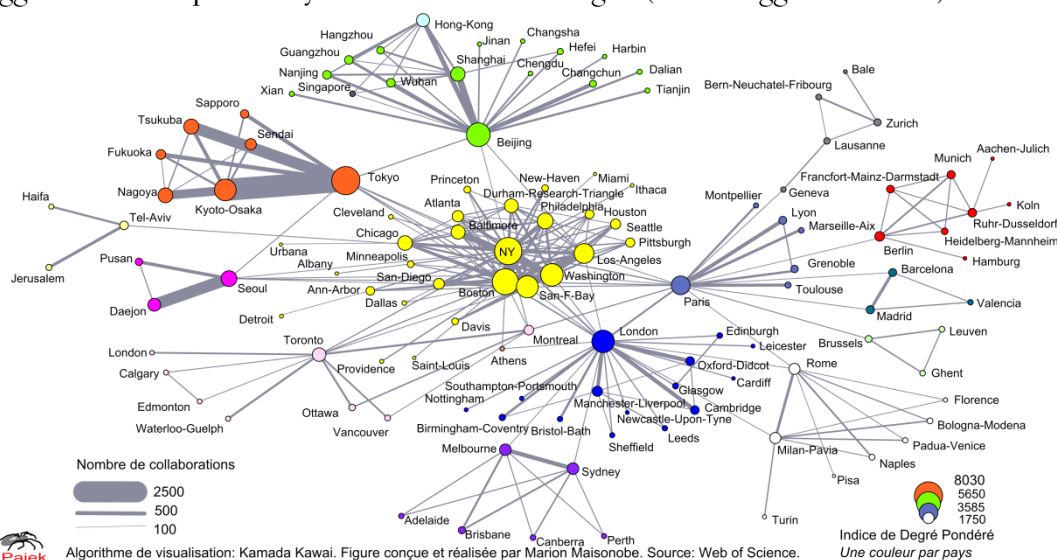


Figure 42 – Composante principale du réseau mondial des collaborations scientifiques en 2007. Source : Web of Science, moyenne mobile sur 3 ans

En-dessous, la représentation des collaborations scientifiques (Figure 42) a été obtenue à partir du graphe des co-signatures d'articles en 2007 (moyenne mobile sur 3 ans). Ne sont conservées que les relations d'une valeur supérieure à 100. Le résultat montre que très peu de relations internationales dépassent ce seuil. En particulier, puisque les agglomérations néerlandaises, russes, suédoises, taiwanaises, polonaises, turques et brésiliennes n'entretiennent aucune relation internationale de cette intensité, elles ne sont pas reliées à la composante principale du réseau représenté ici, c'est-à-dire au plus grand ensemble connexe du graphe. Sur la figure, la couleur des cercles symbolisant les villes dépend de leur appartenance à des pays. Ce partitionnement est intéressant car il met en évidence le caractère national des liens de collaborations scientifiques les plus développés. Il ressort que les collaborations internationales représentent une faible part des collaborations interurbaines. Enfin, on constate assez nettement que ce monde scientifique « visible » s'articule autour des États-Unis. En effet, une part considérable des liens internationaux supérieurs à 100 sont entretenus avec des villes américaines. Réalisée à partir du logiciel d'analyse de réseaux Pajek, cette représentation a été retouchée et légendée à l'aide d'Inkscape.

L'information représentée sur ces images est extrêmement allégée par rapport à l'information de départ. Tant sur la carte que sur le diagramme nœud-lien, il convient de filtrer l'information pour la rendre lisible, mais cela se fait au prix de l'« invisibilisation » du reste du système scientifique mondial. Pour les données relationnelles, d'autres solutions peuvent être envisagées comme l'a montré le géographe Stephen Henneman dans un article stimulant, qui interroge l'opportunité d'avoir recours à la carte, au diagramme nœud-lien ou encore au diagramme circulaire ou de *chord* pour représenter des liens de collaborations scientifiques entre lieux (Hennemann, 2013). Dans le même esprit, pour donner à voir sous différentes formes et de façon plus exhaustive les données relationnelles du WoS, nous nous sommes penchée avec Laurent Jégou sur l'opportunité du recours à des solutions de représentations interactives. En particulier, nous avons cherché ensemble un moyen de valoriser et visualiser le contenu des matrices de lieux obtenues pour plusieurs dates à l'issue du comptage fractionné des co-signatures. Ces matrices valuées sont des objets regroupant une densité d'information telle qu'il est impossible d'en saisir le contenu en un seul coup d'œil. À cet effet, une plateforme a été mise au point (à l'aide de bibliothèques de fonctions *javascript*) qui permet de faire varier les modes de visualisation, les périodes étudiées, et la résolution de l'analyse à partir de ces matrices, à consulter sur le site : <http://www.Coscimo.net/>. En Octobre 2013, ce projet a remporté le troisième prix du concours de géovisualisation et cartographies dynamiques du Festival International de Géographie de Saint-Dié-des-Vosges (Jégou & Maisonobe, 2013).

Ce projet « Coscimo » pour « collaborations scientifiques mondiales » porte essentiellement sur la question de la visualisation ou plutôt de la représentation des données de collaborations scientifiques. Il est parti de l'hypothèse que l'interactivité pouvait être une solution intéressante pour visualiser les données spatialisées de collaborations scientifiques. L'intérêt de Coscimo, au-delà de son contenu, est de trouver des solutions de visualisation qui tiennent compte de la multidimensionalité des données manipulées. L'application s'appuie sur les co-publications indexées dans le WoS toutes disciplines et types de publication confondus. Divers modes de visualisation interactive des collaborations scientifiques sont proposés et l'exploration des résultats est disponible à plusieurs niveaux de résolution (villes, pays, continents) pour deux dates (2000 et 2007, moyenne mobile sur 3 ans). Quatre modes de représentation sont retenus : il s'agit de la représentation matricielle, des diagrammes de *chord* et nœud-lien et de la carte de flux. Les commentaires guidant l'utilisateur montrent que les phénomènes remarquables varient d'un mode de représentation à l'autre. Avec l'interactivité, il est possible d'accéder à davantage d'information qu'à travers un mode de visualisation statique. En revanche, il est parfois difficile de saisir « le message » et de dépasser l'aspect ludique de la navigation, d'où l'accompagnement des commentaires. Enfin, les boutons installés à cet effet permettent de basculer d'une période et d'un niveau d'analyse à l'autre en quelques clics.

D'un point de vue matériel, ce projet a permis d'évaluer l'utilisation de bibliothèques de fonctions *JavaScript* mettant en œuvre les ressources du format HTML5. Ces nouvelles techniques présentent l'intérêt d'être accessibles à la fois pour le public, car le support de la page Internet est quasi universel, mais aussi pour les auteurs, car leur mise en pratique ne nécessite pas de programmation avancée. Les bibliothèques de fonctions utilisées par Coscimo sont : *Data-Driven Documents*, ou D3, développées à l'origine pour les infographies des sites Internet de grands organes de presse comme le *New York Times* ; et SigmaJS, dédiée à la visualisation avancée de graphes. Grâce à ces ressources, le site Internet présente des représentations graphiques originales, légères (car en mode vectoriel), interactives, alimentées par des données fixes ou dynamiques (une base de données) ; tout cela sans nécessiter d'infrastructure serveur lourde (car les représentations sont dessinées par le navigateur Internet de l'utilisateur) (*ibid.*).

Le site ainsi créé se compose de plusieurs pages. Après une présentation générale des données utilisées, leurs sources, leurs caractéristiques spatiales, et l'objectif général des recherches menées (Figure 43), les différentes modalités de représentation graphique sont accessibles depuis les onglets. Elles sont chaque fois accompagnées d'une présentation et d'un court commentaire.

Les collaborations scientifiques mondiales, 1999-2008

Accueil	Présentation	Matrices	Cordes	Cartes	Graphes
---------	--------------	----------	--------	--------	---------

L'article est devenu le mode privilégié de diffusion de la science au niveau mondial. Il est le meilleur indicateur existant pour analyser (localiser et comptabiliser) les collaborations scientifiques. Si le niveau local reste l'aire privilégiée de collaboration du chercheur (environ les deux-tiers), une part non négligeable de ces collaborations est nationale, régionale ou internationale. L'exploitation des données de collaboration scientifique permet de dégager la structure spatiale de la communauté scientifique mondiale. C'est cette dimension supra-locale de l'activité scientifique que nous vous présentons.

La répartition du total des publications selon ce critère des distances de collaboration est présentée dans les graphiques ci-dessous, pour deux périodes récentes.

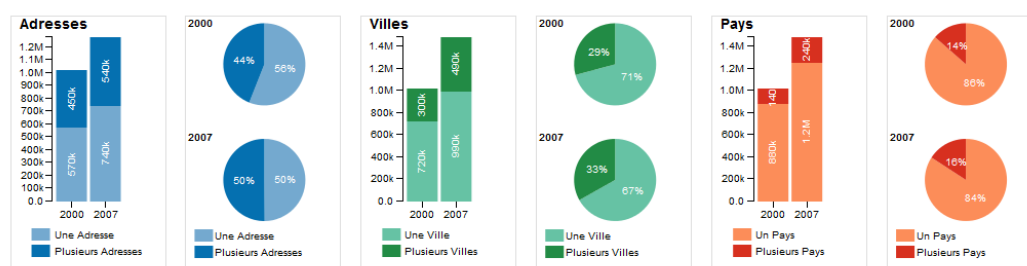


Figure 43 – Présentation générale de la recherche et histogrammes

Le premier mode de représentation est matriciel : comme le défendent Henry et Fekete (Henry & Fekete, 2008), ce type de visualisation classique permet de prendre conscience des vides, c'est-à-dire des relations absentes ou faibles, qu'on a autrement tendance à oublier (Figure 44). En visualisant les matrices, il est aisé de constater que les liens importants sont rares et concernent en priorité les agglomérations et pays appartenant aux mêmes continents (les carrés sont colorés lorsque les relations sont intra-continenteles). Enfin, la matrice est ordonnable avec deux possibilités de tri : par valeur ou continent.

Les collaborations scientifiques mondiales, 1999-2008

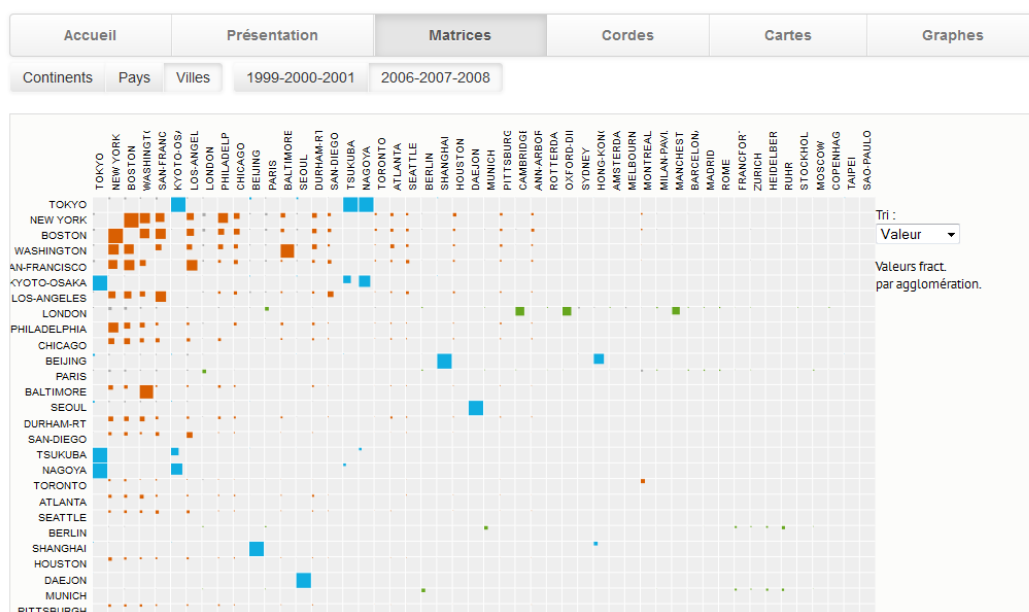


Figure 44 – Matrice des collaborations scientifiques entre agglomérations en 2007

Le deuxième mode de représentation est un diagramme de *chord* : il est plus contraignant en nombre d'entités en relation pouvant être représentées simultanément (ici et au choix, parmi la vingtaine d'entités spatiales les plus publiantes, ces entités sont soit des agglomérations, des pays ou des continents). En plus de sa physionomie séduisante, le diagramme de *chord* permet de rendre compte de l'ordre de grandeur des relations (Figure 45). Venu de la presse d'information, il s'est récemment popularisé en géographie puisque peu de temps après la mise en ligne de Coscimo, des géographes australiens ont proposé, dans la revue *Science*, un mode de représentation identique pour visualiser des flux de population intercontinentaux (Abel & Sander, 2014 ; Carré, 2013). Aussi, le président de l'Université de New-York a mobilisé des *chords* pour une communication généraliste sur l'état des collaborations scientifiques internationales en 2011 à partir d'une « sélection de revues » (Sexton, 2012). En janvier 2015, le *package* R *Circlize* est sorti, qui permet de réaliser de tels diagrammes en appelant la fonction « *chordDiagram* ».

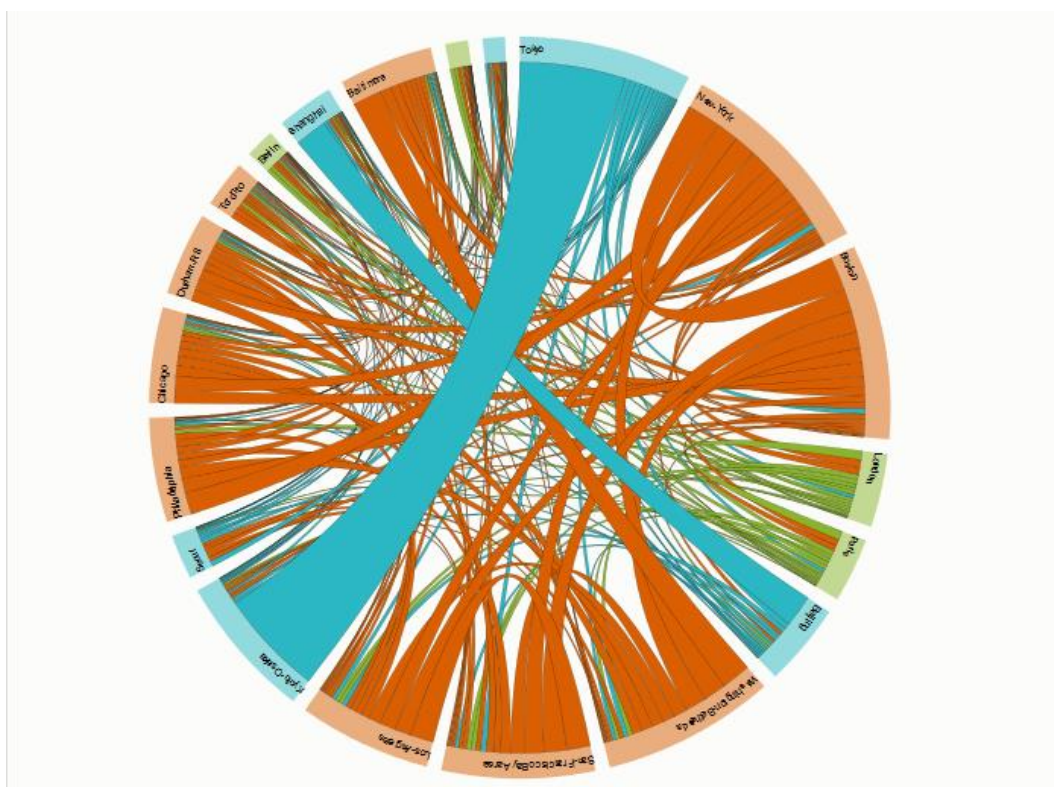


Figure 45 – Diagramme de chord des collaborations scientifiques entre pays en 2007

Dans l'onglet suivant du site Coscimo, une représentation plus classique de type « carte de flux » est proposée (Figure 46). L'intérêt de l'interactivité est ici très important : au lieu de présenter toutes les collaborations existantes entre les 200 premières agglomérations scientifiques mondiales, la carte présente celles d'une agglomération à la fois, ce qui facilite sa lisibilité.

La comparaison entre les profils de collaborations s'effectue dans un deuxième temps, en « balayant » à la souris la carte d'agglomération en agglomération. Ce faisant, la représentation en symboles et en flux proportionnels est recalculée et redessinée en temps réel par le navigateur Internet. Aussi, quand on place son curseur sur une agglomération, la taille du cercle nous indique l'importance du nombre de collaboration de ce lieu et seuls ses liens avec les autres agglomérations du monde apparaissent (l'exemple du réseau égo-centré de Berlin est représenté sur la Figure). Le volume d'information accessible est donc très important par le biais de cette représentation.

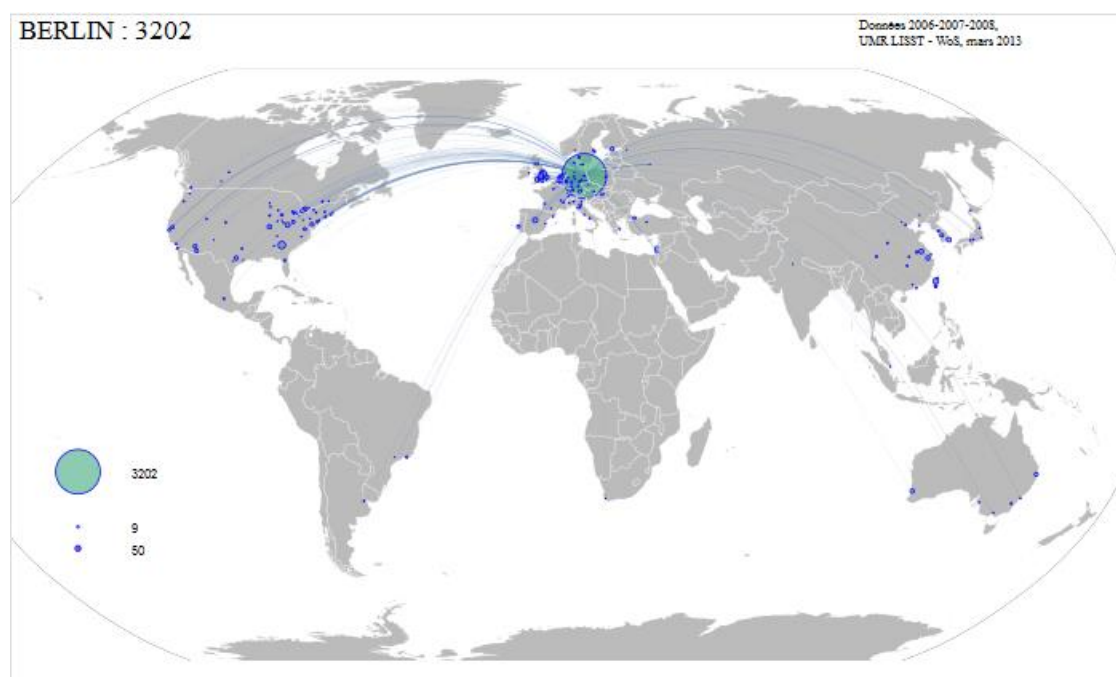


Figure 46 – Carte de flux des collaborations scientifiques entre agglomérations en 2007

Dans la même rubrique du site, l'internaute peut accéder à une carte de synthèse qui donne à voir le résultat d'une analyse de réseaux dont on discutera plus loin (Figure 47). En bref, il s'agit du résultat d'une procédure dite de « détection de communauté » : les villes sont affectées à des communautés ou groupes d'après leur profil relationnel, autrement dit à des ensembles de lieux très reliés entre eux et relativement moins reliés aux autres. Sur la carte, les agglomérations appartenant au même groupe arborent la même couleur. Le résultat montre qu'il existe une adéquation entre le profil relationnel des agglomérations et leur position géographique : on distingue une aire d'influence américaine, une aire d'influence européenne et une aire asiatique. Ici l'interactivité se limite à renseigner le nom de la ville sur laquelle le curseur s'arrête et son groupe d'appartenance.

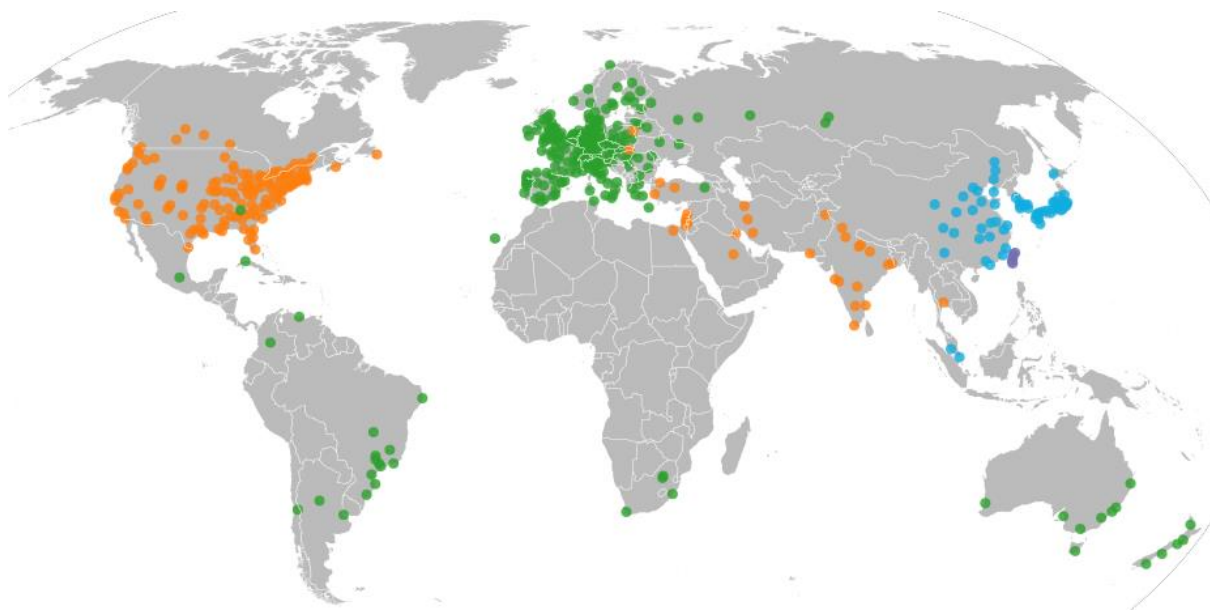


Figure 47 – Résultat du partitionnement de la matrice de collaboration scientifique entre agglomérations en 2007, obtenu en appliquant la méthode de détection de communautés « VoS Clustering », une variante de la méthode de Louvain

Pour finir, le site propose une représentation interactive des collaborations avec le diagramme nœud-lien (Figure 48). Avec ce dernier mode de représentation, c'est la forme de la structure obtenue qui importe. Les collaborations sont seuillées pour ne garder que les liens supérieurs à 200. La position des agglomérations a de l'importance : elle dépend de l'intensité des relations entre toutes les agglomérations représentées (y compris entre celles dont les liens ne sont pas d'une intensité suffisante pour être représentés). Les repères des positions ou coordonnées sont le résultat de l'application d'un algorithme inspiré des méthodes d'analyse multidimensionnelles appelé « VoS Mapping », sur lequel on reviendra. Le fait que la position des agglomérations ne soit pas aléatoire ajoute une richesse à la représentation. L'interactivité permet de zoomer sur différentes parties du réseau mais, en l'occurrence, nous pensons que d'autres avantages de l'interactivité mériteraient d'être testés pour exploiter davantage les possibilités de ce mode de représentation.

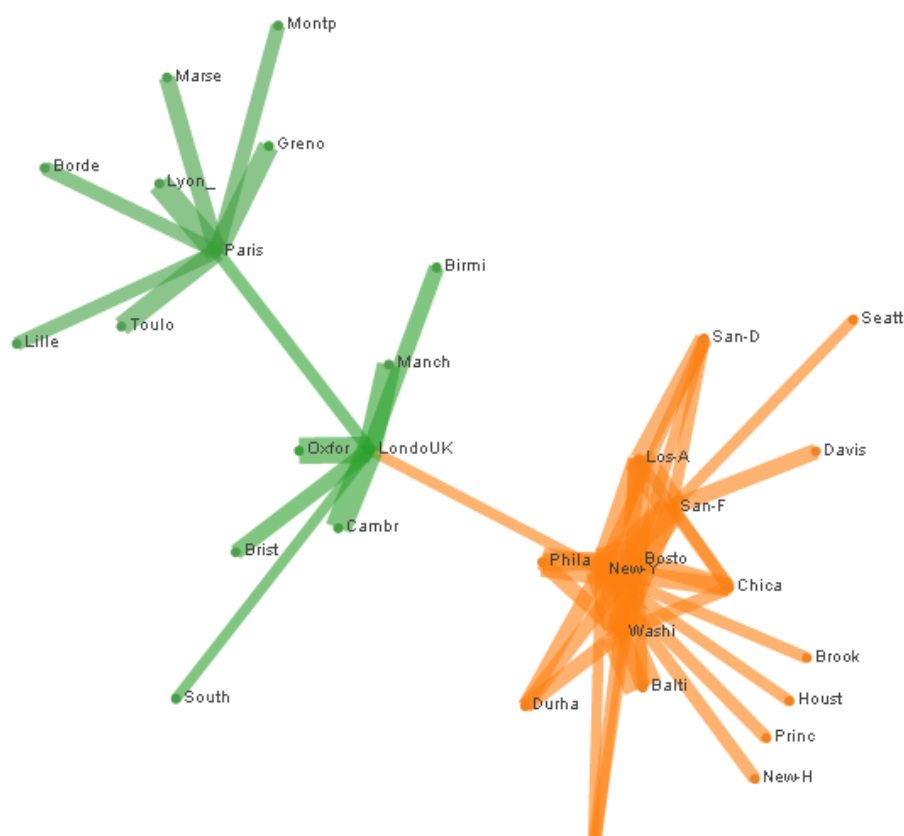


Figure 48 – Graphe des collaborations scientifiques entre agglomérations en 2007

Pour faire le bilan du projet Coscimo, quatre éléments méritent d'être soulignés :

1. D'abord, Coscimo montre l'intérêt de la visualisation des données à l'étape de la communication des résultats d'une recherche. Au bout de la chaîne de traitements des données bibliométriques décrite précédemment, on dispose de matrices évaluées de lieux. Or, le contenu de ces matrices étant trop vaste pour être résumé à l'aide d'une seule image, l'interactivité offre un moyen d'explorer les collaborations scientifiques entre lieux sous différents angles.
2. Mais, cette expérience ne doit pas faire oublier que la visualisation interactive des données peut remplir des fonctions essentielles à d'autres étapes de la chaîne de travail. Pour rappel, l'utilitaire de *wepmapping* développé par Laurent Jégou en amont de nos recherches exploitait les fonctions de l'interactivité pour permettre à l'utilisateur de tester et améliorer la qualité du géocodage, construire des périmètres d'agglomération, explorer les informations bibliographiques aux différents stades de la recherche et ainsi formuler et tester des hypothèses.
3. Cette expérience ne doit pas faire oublier non plus l'intérêt des représentations statiques permettant de résumer efficacement l'information en une seule image. La représentation statique remplit une fonction de synthèse de l'information et peut, à cet égard, être pleinement perçue comme un résultat de recherche.

4. En somme, Coscimo s'accompagne de la conviction que, bien utilisée et bien construite, l'image est en mesure d'être à la fois : outil de recherche, de démonstration ou encore support de communication. C'est pourquoi, il nous importait de guider l'utilisateur de la plateforme tout au long de sa visite, afin qu'il ne reste pas confiné à la dimension ludique du visionnage d'images animées.

Ainsi, le problème posé par la visualisation de données relationnelles spatialisées au niveau mondial a trouvé une solution dans le projet Coscimo. Exploiter les ressources de différents modes de représentations, faire varier les niveaux d'agrégation et les périodes d'analyse, a permis d'arriver à une appréhension multidimensionnelle des collaborations scientifiques mondiales. Mais un grand nombre d'options restent encore à explorer dans cette perspective.

À vrai dire, ce projet arrive dans un contexte où l'exploitation de solutions interactives pour la visualisation des données de recherche en sciences humaines et sociales et leur communication en ligne ont pris une importance croissante. On parle même d'un nouveau domaine de recherche, les « humanités numériques ». Ces dernières mettent en avant et invitent à approfondir la maîtrise des méthodes de traitement, d'exploration, d'analyse, de visualisation et de communication des données numérisées. Regardant l'étude des sciences, une initiative entend illustrer le potentiel de ce champ de recherche, Palladio⁷⁴. Cette plateforme relève du projet : « *Networks in History* » repéré à l'occasion d'un travail de veille réalisé avec Laurent Jégou pour les besoins du portail de recherche auquel nous travaillons : « Géoscimo ». À l'origine de Palladio, il y a un projet de recherche plus classique piloté par l'université de Stanford et intitulé « *Mapping the Republic of Letters* » qui, à travers plusieurs études de cas, a permis de réfléchir aux méthodes de visualisation des données historiques sur les réseaux scientifiques et les réseaux de correspondance en particulier⁷⁵. Fondé sur cette expérience, le projet Palladio vise à rendre accessible en ligne, à tous ceux qui pourraient en avoir l'usage, les solutions de visualisation développées dans le cadre du projet « *Mapping the Republic of Letters* ». En chargeant des données sur cette interface, l'utilisateur est libre d'y appliquer des filtres temporels et d'en explorer les différentes dimensions selon les modalités prévues par les développeurs.

La production de représentations graphiques statiques et interactives arrivant à l'issue du processus de traitement de l'information bibliographique ne sont cependant qu'une des facettes du travail de recherche. Toute cette information sur les lieux de production et les réseaux de collaborations scientifiques, il faut surtout pouvoir en déduire un savoir ou discours géographique à travers une opération de description.

⁷⁴ Source : URL : <http://palladio.designhumanities.org/#/>

⁷⁵ Source : URL : <http://republicofletters.stanford.edu/>

Or, la visualisation n'est pas le seul moyen sur lequel nous appuyer pour décrire les observations. Il existe d'autres méthodes pouvant se combiner à la visualisation. En particulier, puisque nous disposons d'observations quantitatives à plusieurs dates : il est utile de raisonner en relatif, de mesurer des évolutions et d'en déduire des tendances. Cela suppose d'avoir recours aux outils de la statistique. Au-delà, les outils de l'analyse de réseaux permettent de décrypter la structure des réseaux de collaborations scientifiques.

C'est seulement après avoir considéré et pris la mesure de l'apport des différentes méthodes d'analyse, en particulier la statistique descriptive et l'analyse de réseaux combinées aux méthodes de visualisation, que nous les mobiliserons à l'appui de la construction, à partir de nos observations, d'un discours argumenté et suffisamment fondé empiriquement pour ne pas pouvoir être taxé d'idéologique sur la géographie des activités scientifiques contemporaines et ses récentes évolutions. Avant d'en arriver à la production de ce discours, il convient de convoquer les méthodes classiques de l'analyse spatiale, revenir sur les différentes approches de l'espace qu'elles autorisent et, ainsi, justifier le recours privilégié à une approche de l'espace comme support et en relatif pour l'analyse de la production et des collaborations scientifiques à l'échelle mondiale.

Dans cette dernière section méthodologique, chaque approche est mise en relation avec des exemples d'application aux données sur la science, accompagnés de renseignements sur la mise en œuvre technique des méthodes évoquées. Quelques réflexions sont également proposées sur les logiciels mobilisés et leur logique d'utilisation. Si des résultats quantitatifs apparaissent d'ores et déjà dans cette section, il n'est pas encore question d'en discuter dans le détail. Pour pouvoir tirer de ces résultats des interprétations quant à la géographie contemporaine des activités scientifiques et ses dynamiques, il faudra procéder à un travail de contextualisation des résultats obtenus, nous permettant de déterminer s'ils s'inscrivent dans la continuité ou la rupture des connaissances et des idéologies circulant au sujet de cette géographie. Le Chapitre 9, venant clôturer cette seconde partie de la thèse, centrée sur le travail d'élaboration d'une méthode de scientométrie spatiale en équipe et d'obtention de résultats généraux à l'échelle mondiale, sera le lieu de cet effort de contextualisation, de restitution et d'interprétations des grandes tendances contemporaines que nos données ont permis de mesurer sur la géographie des activités scientifiques dans le monde, toutes disciplines confondues.

2. Les méthodes d'analyse spatiale quantitative et de réseaux

« Entre les composantes élémentaires d'un même ensemble territorial, articulées les unes aux autres, existent habituellement des différences ou des inégalités. Au fil du temps, ces inégalités se renforcent, s'atténuent ou s'inversent selon la résultante des flux d'hommes, d'argent, de marchandises et d'information, de telle sorte qu'un sous-ensemble territorial se définit moins par ses caractères intrinsèques que par les liens qu'il entretient avec des sous-ensembles territoriaux appartenant au même degré de l'échelle. » (Reynaud, 1992, p. 21-22)

En géographie, l'analyse spatiale est la pratique de la discipline permettant d'interroger des « tableaux d'information géographique ». Ces tableaux présentent une spécificité : « Par rapport aux simples tableaux statistiques, les tableaux d'information géographique ont la particularité de décrire des entités spatiales, c'est-à-dire des individus statistiques qui ne sont pas *a priori* indépendants les uns des autres, mais qui sont au contraire pris dans des relations de voisinage, de position relative, ou d'emboîtement. » (Pumain & Saint-Julien, 1997, p. 17). Le traitement de l'information géographique des notices bibliographiques permet d'aboutir à ce type de tableaux. En effet, une fois l'information agrégée par entités spatiales à l'aide des méthodes de comptage précédemment définies, on peut en déduire deux types de tableaux :

1. Un tableau indiquant un nombre de publications par entités spatiales avec les entités spatiales en ligne et l'année de publication en colonne. Lorsqu'on travaille au niveau des « agglomérations scientifiques », ce tableau peut comporter plus de 10 000 lignes (une ligne par agglomération scientifique publiante). La figure ci-dessous (Figure 49) montre un extrait d'un tableau de ce type, obtenu à partir de l'ensemble des articles, recensions et lettres du SCI Expanded. Les chiffres donnent le nombre de publications par année par agglomération. Les deux dernières colonnes donnent les chiffres en moyenne mobile sur 3 ans pour les deux séries d'années considérées (1999-2001 ; 2006-2008).

nom	pays	idcomposite	af_sd_1999	af_sd_2000	af_sd_2001	af_sd_2006	af_sd_2007	af_sd_2008	af_sd_9901	af_sd_0608
TOKYO	JAPAN	AD#7725	20662,639	21310,504	21690,856	24151,403	25020,361	25189,137	21221,333	24786,967
BEIJING	CHINA	AD#2074	5627,7247	6601,8537	8180,3423	18151,925	20453,163	22577,455	6803,3069	20394,181
PARIS	FRANCE	AD#3612	14617,342	14383,095	14363,377	15460,418	16158,846	16966,521	14454,605	16195,261
NEW YORK	UNITED-STA	AD#15377	13481,077	13998,534	14271,396	15874,111	16138,122	16248,823	13917,002	16087,019
SEOUL	SOUTH-KORE	AD#9989	5205,2608	6063,6891	7218,2179	14288,413	16387,056	16902,923	6162,3892	15859,464
BOSTON	UNITED-STA	AD#12860	12019,416	12076,82	12594,499	14948,486	15041,562	15598,705	12230,245	15196,251
LONDON	UNITED-KING	AD#11950	13976,192	14274,607	14028,399	14797,625	14962,86	15044,834	14093,066	14935,107
KYOTO-OSAKA	JAPAN	AD#7568	13227,872	13003,756	13045,694	14196,153	14654,327	14072,973	13092,441	14307,818
SAN FRANCISCO	UNITED-STA	AD#16121	10576,558	10491,139	10921,031	12645,976	12930,195	12957,512	10662,909	12844,561
WASHINGTON	UNITED-STA	AD#16876	10500,377	10778,775	11222,935	12421,903	12385,891	12432,021	10834,029	12413,272
LOS ANGELES	UNITED-STA	AD#15644	8741,7889	8961,1848	9566,6118	12168,259	12206,499	12425,48	9089,8618	12266,746
SHANGHAI	CHINA	AD#2274	2459,4638	2892,8884	3485,6684	8705,5031	9882,0401	11200,615	2946,0069	9929,3861
MOSCOW	RUSSIA	AD#9414	9924,2543	9794,8746	9110,884	8651,8514	9406,4903	9728,4261	9610,0043	9262,2559
TAIPEI	TAIWAN	AD#10900	4291,7953	4465,3373	5160,9826	8110,3645	8728,3923	9552,6112	4639,3717	8797,1227
CHICAGO	UNITED-STA	AD#13188	6431,3993	6381,1913	6634,3198	7642,5744	8152,1946	7769,0413	6482,3034	7854,6034
PHILADELPHIA	UNITED-STA	AD#15709	6421,5924	6442,6422	6722,7635	7314,6569	7400,9247	7793,5158	6528,9994	7503,0325

Figure 49 – Extrait de la table indiquant le nombre de publications par agglomération par année. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

2. Un tableau d'échanges ou matrice d'adjacence qui s'obtient à partir de l'ensemble des publications co-signées par au moins deux entités spatiales élémentaires indiquant un nombre de collaborations par couple d'entités spatiales. La valeur de la collaboration entre deux entités spatiales se trouve au croisement entre la ligne et la colonne correspondant respectivement à l'une et l'autre. Quand les données sont agrégées au niveau des agglomérations scientifiques, ce tableau peut alors aller jusqu'à décrire près de 50 millions de relations singulières (sur près de 100 000 millions de relations renseignées dans le tableau par année). De fait, puisque la matrice est symétrique : il n'y a que la moitié du tableau qui importe (l'autre moitié est une duplication de la première). En témoigne l'image ci-dessous (Figure 50) extraite de la matrice des collaborations entre agglomérations en 2007 (moyenne mobile sur 3 ans). Comme on ne comptabilise pas les collaborations intra-agglomération, la diagonale de la matrice est vide lorsqu'on travaille au niveau des agglomérations. Il en va autrement si l'on considère les niveaux d'agrégation ou de résolution supérieurs à cette entité de base (niveau national ou continental par exemple).

		1	2	3	4	5	6	7	8	9	10
1 AD#7725	"Tokyo_Japar	0	139,323822	104,41798	76,5941443	82,9577736	104,797105	125,139555	2416,60615	88,7387098	123,026479
2 AD#2074	"Beijing_Chir	139,323822	0	91,4384953	64,4787861	49,8301455	68,774269	49,7475274	73,7284195	80,2590114	84,6260433
3 AD#15377	"New-York_U	104,41798	91,4384953	0	131,84444	126,96084	581,224062	84,5487536	52,4100906	395,293733	370,111249
4 AD#11950	"London_Unit	76,5941443	64,4787861	131,84444	0	175,009225	119,962929	21,0582194	39,1050055	90,3242912	83,7337954
5 AD#3612	"Paris_France	82,9577736	49,8301455	126,96084	175,009225	0	114,486375	26,0058681	40,8418772	81,8369522	105,754396
6 AD#12860	"Boston_Unit	104,797105	68,774269	581,224062	119,962929	114,486375	0	102,473944	57,5448362	397,221622	414,884151
7 AD#9989	"Seoul_South	125,139555	49,7475274	84,5487536	21,0582194	26,0058681	102,473944	0	73,8277956	91,3760715	89,0308809
8 AD#7568	"Kyoto-Osaka	2416,60615	73,7284195	52,4100906	39,1050055	40,8418772	57,5448362	73,8277956	0	38,9435588	42,4659825
9 AD#16876	"Washington	88,7387098	80,2590114	395,293733	90,3242912	81,8369522	397,221622	91,3760715	38,9435588	0	252,49055
10 AD#16121	"San-Francisc	123,026479	84,6260433	370,111249	83,7337954	105,754396	414,884151	89,0308809	42,4659825	252,49055	0
11 AD#15644	"Los-Angeles	87,1344544	103,3361	283,492847	68,6317221	88,224659	306,206541	76,8024045	44,1340198	228,617296	450,76991
12 AD#2274	"Shanghai_Ci	69,5343361	835,158908	49,8636797	27,2352107	41,9189705	28,2404514	19,4094041	58,8541599	24,2241396	43,999922
13 AD#9414	"Moscow_Ru	38,7253798	19,8776692	39,702309	32,6168924	93,851625	30,2228284	11,9162933	19,995873	32,8539798	16,8762448
14 AD#13188	"Chicago_Un	25,8154659	49,6059665	248,075812	39,0030334	51,6713709	222,186278	52,6243278	23,598996	179,202292	179,511492
15 AD#10900	"Taipei_Taiw	62,769639	46,4936821	41,1538198	22,3907785	9,1829707	43,5138301	9,71290538	19,4161042	25,0253606	38,0182669
16 AD#15709	"Philadelphia	29,1556156	41,2852195	362,365179	65,4405175	42,2566179	244,833533	24,9404958	17,4168756	203,443497	128,706864
17 AD#1975	"Toronto_Car	41,8286229	41,3708729	100,461523	58,6164172	38,0526854	110,674897	33,6722722	18,2999042	57,2538574	85,3789025
18 AD#13561	"Durham-Res	20,5263718	29,8613783	205,010551	46,4068654	25,6180273	191,131366	54,9340438	24,5577958	211,083757	130,627772
19 AD#10167	"Madrid_Spai	9,19620135	6,96093371	55,8080762	68,2635108	94,8316639	31,0078501	4,48104421	5,90462953	25,8222664	27,7857385
20 AD#709	"Sydney_Ausi	24,233211	49,1549243	35,5624123	92,0436683	35,1182744	41,8378876	14,8618804	12,6569851	33,8854041	27,5035978
21 AD#4076	"Berlin_Germ	29,6993999	41,3175289	54,9221417	73,3847966	86,1829011	51,3978582	10,3718139	11,5769949	21,4510326	45,9312631
22 AD#5272	"Hong-Kong	21,8457525	552,492563	38,5553655	61,5201761	20,1432203	30,0782121	14,4394996	10,999664	17,319136	28,0542224

Figure 50 – Extrait de la matrice de collaborations scientifiques entre agglomérations à l'année 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Pour résumer, exploiter et analyser le contenu de ces tableaux, il existe un large éventail de possibilités. Les méthodes à convoquer dépendent évidemment des questions posées. Dans le présent chapitre, nous allons discuter des « boîtes à outil » dans lesquels piocher pour analyser la géographie de la production scientifique, et en particulier la statistique descriptive et l'analyse de réseaux. Ces techniques sont précieuses pour décrire la répartition des entités spatiales et l'évolution de cette répartition d'une part ; le système relationnel des entités spatiales et l'évolution de la structure de ce système d'autre part.

La statistique descriptive est utilisée dans un grand nombre de disciplines et à l'appui de nombreuses méthodes d'analyse pour décrire un phénomène à partir d'observations chiffrées, résumer l'information à partir de grands ensembles de données, raisonner en valeurs relatives à l'aide de pourcentages. C'est donc un outil fondamental pour faire de la scientométrie. Comme la visualisation, la statistique descriptive représente les données. Mais, en général, les approches classiques de la statistique descriptive ne satisfont pas entièrement les géographes puisque : « les méthodes de la statistique descriptive (...) traitent seulement de la partie sémantique de l'information géographique, la partie géographique n'étant utilisée qu'*a posteriori*, pour faire une représentation cartographique des résultats » (Pumain & Saint-Julien, 1997, p. 17). Par exemple, après avoir mesuré des évolutions ou avoir effectué une classification, il est parfois bienvenu d'en cartographier les résultats. Mais cette utilisation des statistiques en analyse spatiale reste limitée. Dans une démarche de scientométrie spatiale, il paraît également important de réfléchir à l'opportunité de combiner une approche géomatique (pour tenir compte de la composante géométrique de l'information) à une approche statistique.

Aussi, dans la mesure où le « tableau d'échange entre unités spatiales » n'est rien d'autre qu'une « matrice de flux », il est parallèlement tout à fait possible d'en tirer des graphes et de le soumettre aux méthodes de l'analyse de réseaux. Dans le manuel d'analyse des réseaux sociaux de Pierre Mercklé, on trouve que pour étudier ce type d'objets, « l'apport méthodologique de la théorie des graphes est double : d'une part les graphes donnent une représentation graphique des réseaux de relations, qui facilite leur visualisation, permet la mise en lumière d'un certain nombre de leurs propriétés structurales ; d'autre part, la théorie des graphes développe un corpus extrêmement riche de concepts formels permettant de mesurer un certain nombre de propriétés des relations entre éléments. » (Mercklé, 2004, p. 25). En effet, outre les possibilités de visualisation évoquées dans la section précédente, la théorie des graphes offre « un arsenal de concepts, de théorèmes, d'algorithmes et de raisonnements, grâce auxquels la figuration graphique dépasse la simple représentation pour autoriser un traitement mathématique générateur de connaissances nouvelles » et c'est ce qui « constitue son second apport fondamental à l'analyse des réseaux. » (*loc. cit.*). Parmi les outils les plus utilisés de ce domaine d'analyse, on trouve les algorithmes qui servent à la recherche de sous-groupe constitutifs d'une structure, les indices locaux qui permettent de caractériser les nœuds ou les liens ou encore les indices mettant en évidence les propriétés globales du graphe. En s'inscrivant dans la continuité des travaux pionniers en scientométrie spatiale, nous montrons que les méthodes de l'analyse des réseaux méritent d'être utilisées pour étudier la géographie des collaborations scientifiques.

À l'issue de cette dernière section méthodologique, nous aurons montré comment la statistique descriptive et l'analyse de réseaux permettent de saisir et d'expliquer la répartition de la production scientifique et son organisation spatiale, le tout en dynamique.

2.1. La répartition spatiale des activités scientifiques

Comme en témoignent les résultats présentés par la suite, notre méthode calibrée pour traiter l'ensemble du *Web of Science* (WoS) peut être déclinée pour traiter des sous-ensembles du WoS comme les trois grands index (SCI Expanded, SSCI et AHCI), de grands domaines disciplinaires ou de plus petits corpus, voire même, on le verra, des corpus bibliographiques extraits d'autres bases de données comme *Scopus*. En attendant, indépendamment de connaissances préalables ou d'hypothèses précises sur l'ensemble de données à analyser, il s'agit de discuter des outils d'analyse et des logiciels à envisager une fois munis d'un ensemble de données bibliographiques géolocalisées et agrégées au niveau d'entités spatiales.

Dans ce cas, il est commun de commencer par la question « Où », c'est-à-dire par l'étude de la répartition géographique du phénomène représenté par le corpus. Pour saisir cette répartition, on distingue généralement deux approches :

- l'étude du semis de points, autrement dit la gestion de l'information concernant la présence ou l'absence de publications par lieu et par année ;
- l'étude de la distribution des volumes de publications par lieu et par année. Cette distribution statistique suit *a priori* une loi de Lotka ce qui veut dire qu'une poignée de lieux concentrent l'essentiel des publications (Encadré 2).

Un différentiel, de simples taux d'évolution, des écarts de valeurs relatives permettent de comparer entre elles plusieurs formes de semis de points et plusieurs distributions statistiques de publications. La comparaison peut se faire entre plusieurs années pour un même corpus, ce qui revient à mesurer une dynamique, ou bien entre plusieurs corpus pour une même année, ce qui revient à comparer différents phénomènes concomitants. Évidemment, on peut se concentrer sur la répartition de sous-ensembles du corpus, par exemple les co-publications. On peut aussi agréger ou décrire la distribution de l'information à des niveaux géographiques supérieurs. Et reporter sur des cartes, tableaux ou graphiques toutes les informations quantitatives ainsi obtenues.

Très vite, lorsqu'on se lance dans la description des observations recueillies pour répondre à la question « Où », on est ramené à la question « Où comment ? » (Pumain & Saint-Julien, 1997). En général, cette question invite à déduire un degré de concentration spatiale de la répartition observée. En effet, une telle question suppose de situer la répartition du phénomène entre différents modèles idéaux de répartition spatiale comme la répartition aléatoire, régulière ou concentrée. Là encore, il faut distinguer l'étude de la forme du semis de points et l'étude de la distribution statistique de l'information (de son volume). En cela, il s'agit de distinguer « concentration statistique » et « concentration spatiale » de l'information. Les indices statistiques classiques de concentration, comme l'indice de Gini, n'intègrent pas la notion de dépendance entre les entités d'analyse. En particulier, ils s'occupent de déterminer si les publications sont réparties équitablement entre les différents lieux sans se préoccuper de leur position géographique respective. Or, pour être rigoureuse, l'analyse spatiale doit tenir compte des interdépendances entre les entités d'analyse, parmi lesquelles la distance qui les sépare. C'est là toute la différence entre une approche qui tient compte de la position spatiale des observations et une approche statistique basique qui suppose leur indépendance. Dans un article déjà cité sur l'évolution des périmètres d'agglomérations dans le temps, Philippe Julien attire justement l'attention sur cette subtilité (Encadré 3).

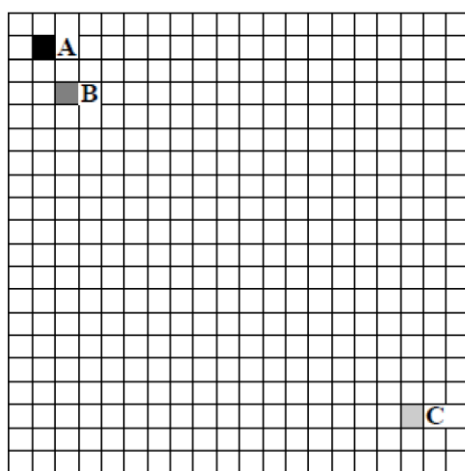
DISTINGUER CONCENTRATION STATISTIQUE ET CONCENTRATION SPATIALE

Pour évaluer la concentration spatiale de la population, les outils classiques du statisticien (du type indice de Gini, de Theil, etc.), utilisés de façon abrupte, sont peu efficaces, justement parce qu'ils sont par nature « a-spatiaux » : ils permettent seulement de calculer une concentration statistique, mais ne tiennent pas compte des dispositions spatiales respectives des objets territoriaux que le calcul agrège.

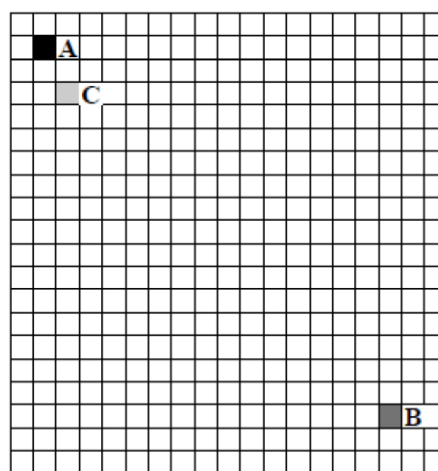
Supposons qu'on ait trouvé une maille territoriale pertinente pour le calcul de la concentration statistique. Toute la dimension spatiale n'aura pas été prise pour autant, comme on peut le montrer sur le schéma théorique ci-dessous.

Prenons deux pays P1 et P2 aux mailles territoriales M à superficies rigoureusement identiques (M désigne la maille territoriale jugée pertinente pour analyser la

concentration). Chacun des 2 pays a 3 mailles habitées, A (1 million d'habitants), B (100 000 habitants) et C (1 000 habitants), les autres mailles ne comptant qu'un habitant : la distribution de population étant identique (1 million, 100 000, 1000, 1,1,1, ...), les deux pays ont exactement la même concentration statistique. Mais la configuration spatiale diffère : dans P1, les villes A et B sont proches, et éloignées de C ; tandis que dans P2, A et C sont proches, et éloignées de B. Le sens commun fait dire que la concentration spatiale de P1 est supérieure à celle de P2. P1 et P2 peuvent représenter aussi le même pays à deux dates différentes si les chiffres de populations sont naturellement plus proches. À l'évidence, une notion de distance doit être introduite, afin de décrire, dans une démarche beaucoup plus complexe, la concentration spatiale.



Pays P1



Pays P2

La couleur noire représente 1 million d'habitants, la grise foncée 100 000 habitants et la grise claire 1000 habitants ; la couleur blanche représente 1 habitant.

Encadré 3 - Distinguer concentration statistique et concentration spatiale. Encadré extrait de l'article « Mesurer un univers urbain en expansion » (Julien, 2000)

Dès lors, comment peut-on s'y prendre pour éviter les mauvaises interprétations quant à la distribution statistique d'un phénomène spatialisé ?

Pour commencer, il suffit d'être sensible aux jeux d'échelle. Par exemple, prenons une étude au niveau des municipalités françaises plutôt qu'au niveau supérieur des agglomérations. Si les communes périphériques de Paris gagnent des publications alors que Paris en perd, les indices statistiques signaleront une déconcentration de la production française au détriment de la capitale ; y compris s'il n'y a que les communes limitrophes de Paris qui ont gagné des publications quand le reste des communes françaises n'a pas progressé (c'est ce que démontre l'Encadré 3). La solution proposée dans l'exemple présent consiste à considérer ce qu'il se passe à un niveau supérieur d'analyse : la région.

En remarquant que la croissance des publications a profité à la région capitale au détriment du reste de la France, l'analyse se révèle plus pertinente qu'en se limitant à interpréter la situation au niveau des municipalités. C'est la leçon véhiculée par le MAUP (*Modifiable Areal Unit Problem*) : le même évènement admet différentes interprétations suivant le niveau d'analyse qui est valorisé. C'est ce qui justifie le souci de combiner les approches et de raisonner à plusieurs niveaux⁷⁶ (Openshaw, 1984 cité par Grasland & Madelin, 2006).

Plus généralement, il convient de tenir compte de l'interdépendance entre entités spatiales, leurs relations d'emboîtement mais aussi de distance ou proximité comme variables dans l'analyse. Pour exprimer la distance, on peut s'en remettre à une variable qualitative ou quantitative, et se placer dans un repère absolu ou relatif.

Il existe, en effet, plusieurs « façons de mettre l'espace en variable » d'après Joël Charre qui en distingue trois (Charre, 1995, p. 53) :

1. à travers des coordonnées (x, y, voire z) ;
2. à travers des qualités, en distinguant des sous-espaces, séparés par des « discontinuités » (cette solution rejoint l'idée précédente selon laquelle il faut considérer les emboîtements ou effets d'appartenance et les effets d'échelle) ;
3. et enfin, « au lieu d'être un référentiel (premier cas) ou une distribution de modalités qualitatives (deuxième cas) », l'espace peut-être vu de manière relative.

Selon l'approche que l'on privilégie, les outils d'analyse à mobiliser diffèrent. Prenons pour les deux premières approches un ou deux exemples de traitements permettant de déduire l'évolution de la concentration spatiale des données de publications. Parce qu'elle justifie le recours à un ensemble de méthodes bien spécifiques (analyses de réseau), l'intérêt de la troisième approche sera évoquée ultérieurement (Section 2).

Pour aborder l'espace comme un référentiel au sein duquel les entités spatiales sont positionnées d'après leurs coordonnées géographiques (latitude et longitude), on mobilise un système d'information géographique (SIG) comme Quantum Gis (QGIS)⁷⁷.

⁷⁶ Le géographe britannique Stan Openshaw est resté associé à cette difficulté en consacrant plusieurs études et travaux au problème d'instabilité des entités spatiales et aux biais statistiques qui en découlent, problème baptisé « *Modifiable Areal Unit Problem* » (1984). Aujourd'hui, ce problème continue de préoccuper les géographes et les pouvoirs publics, en particulier pour produire des analyses robustes à l'échelle européenne, en témoigne un rapport daté de 2006 issu d'une recherche coordonnée par Claude Grasland et Malika Madelin menée dans le cadre du programme European Spatial Planning Observation Network (ESPON) 2000-2006. Ce rapport insiste sur les nombreuses opportunités offertes par le développement des Systèmes d'Information Géographique pour surmonter le MAUP. Il démontre par ailleurs que le MAUP n'est pas un problème en soi. Le simple fait de connaître les risques associés au MAUP permet de se prémunir contre les mauvaises interprétations en analyse spatiale (Openshaw, 1984 cité par Grasland & Madelin, 2006).

⁷⁷ Source : URL : <http://www.qgis.org/fr/site/>

S'il est commode d'utiliser QGIS pour en extraire des représentations cartographiques, il s'agit d'abord d'un SIG, ce qui veut dire qu'il structure l'information d'une manière qui se prête à la formulation de requêtes spatiales. Depuis l'interface graphique de ce SIG, plusieurs opérations d'analyse spatiale sont exécutables qui tiennent compte de la position des entités d'analyse. Nous avons retenu deux exemples d'analyses basiques : la méthode du plus proche voisin et le calcul des coordonnées moyennes d'un semis de points.

Pour étudier la forme d'un semis de points, QGIS dispose d'un module d'« Analyse du plus proche voisin ». Après avoir chargé une couche de points au format *shapefile* (« .shp ») dont on souhaite étudier la répartition, on accède à ce module en allant à : **Vecteur/Outils d'analyse/Analyse du plus proche voisin**⁷⁸

Cette méthode permet de déterminer dans quelle mesure la répartition spatiale d'un semis de points s'éloigne d'une répartition aléatoire. Plus l'indice de dispersion obtenu s'éloigne de 1 – c'est-à-dire plus la moyenne des distances observées au plus proche voisin diffère de la distance théorique moyenne au plus proche voisin dans le cas d'une distribution aléatoire – plus le semis de points s'éloigne d'une répartition aléatoire. S'il se rapproche de zéro l'indice de dispersion indique que le phénomène est concentré dans un nombre limité de zone faiblement étendues sur la surface considérée (Figure 51). On trouve une description complète de cette méthode dans les supports de cours de Claude Grasland, disponibles en ligne⁷⁹ (Grasland, 2000).

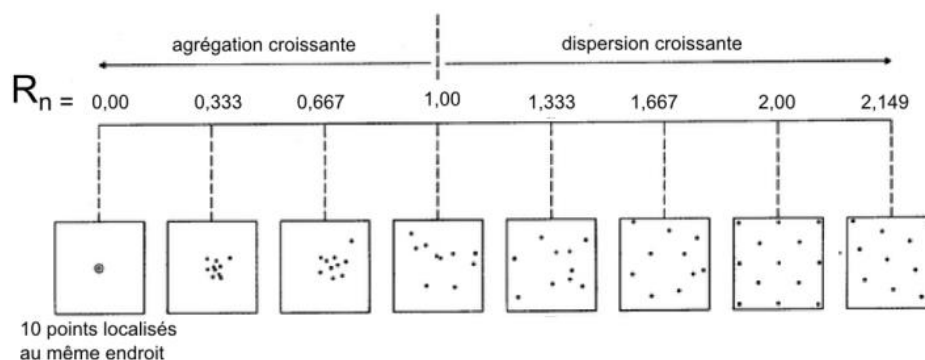


Figure 51 – Forme de la distribution d'après la valeur de l'indice de dispersion (R_n). Figure extraite et traduite du support de cours de Claude Grasland (Grasland, 2000)

⁷⁸ Une console s'ouvre dans laquelle il faut indiquer sur quelle couche de points réaliser l'analyse. Le logiciel renvoie une série de résultats dont on trouve l'explication sur le forum Georezo.net dans un commentaire posté le 30 septembre 2013 et accessible à l'adresse suivante :

<http://georezo.net/forum/viewtopic.php?id=87466>

⁷⁹ Source : URL : <http://grasland.script.univ-paris-diderot.fr/>

Dans l'exemple de la répartition des points de publication scientifique à l'échelle mondiale en 2006-2008, on obtient des valeurs (proches de 0,3) très éloignées de celles qu'on aurait obtenues dans le cas d'une répartition aléatoire. Il faut dire qu'à l'échelle mondiale, l'espace considéré est loin d'être continu contrairement à ce que suppose le modèle. Surtout, une vaste portion de l'espace terrestre est recouverte de mers et d'océans, ce qui rend absurde l'hypothèse de répartition aléatoire des points sur la surface considérée. Enfin, compte tenu de la rotondité du monde, l'utilisation d'une métrique euclidienne revient à considérer une distance qui s'apparente peu à la distance « réelle » ou « géographique » entre les entités spatiales (Grasland, *ibid.*). Néanmoins, quelques déductions peuvent se faire en comparant les valeurs de l'indice de dispersion obtenues pour différentes distributions. Pour plus de commodité, la terminologie « ville » est utilisée pour désigner les entités d'analyse même s'il s'agit en l'espèce d'« agglomérations » de localités publiantes. Ainsi, l'indice de répartition est stable entre 2000 et 2007 pour le semis de villes dont au moins une participation a été enregistrée dans l'ensemble du WoS (toutes publications confondues) alors qu'il tend à diminuer pour le sous-ensemble des villes actives dans le SCI Expanded (articles, recensions et lettres), suggérant une concentration (Tableau 14). On remarque qu'aux deux dates la moyenne des distances observées est plus grande pour le sous-ensemble des villes du SCI Expanded que pour l'ensemble des villes du WoS. Il apparaît cependant que les valeurs moyennes se rejoignent en 2007. Cela provient du fait que le nombre de villes participant au SCI Expanded augmente bien plus entre 2000 et 2007 (+ 2300 villes) que le nombre total de villes dont l'activité scientifique est repérée dans le WoS (+ 1500). Ce faisant, les deux semis de point comparés se ressemblent davantage en 2007 qu'ils ne se ressemblaient en 2000. Or, l'augmentation du nombre de villes se traduit à l'échelle mondiale par un indice de dispersion plus faible car les villes dont l'activité scientifique devient visible ont naturellement tendance à apparaître à proximité de lieux d'activité déjà visibles plutôt qu'au milieu d'une étendue déserte ou océanique ! Ce phénomène accentue l'effet de concentration.

Analyse du plus proche voisin	Villes publiantes (Wos)		Villes publiantes (SCI Exp)	
	2000*	2007*	2000*	2007*
Distance moyenne observée	40196,7	36331,4	45171,3	37865,3
Distance moyenne attendue	117256,4	107673,0	126957,9	110376,7
Index du plus proche voisin	0,34	0,34	0,36	0,34
Nombre de villes	8254	9795	7828	9321
Score Z	-114,2	-125,4	-103,4	-121,3
Part dans la production	100%	100%	100%	100%

*Comptage fractionné (WNC), moyenne mobile sur 3 ans.

Source: WoS (toutes publications) et SCIExp (articles, recensions et lettres)

Tableau 14 – Analyse du plus proche voisin pour l'ensemble des villes publiantes

Le Tableau 15 montre les résultats de l'application de la même méthode à des sous-ensembles de villes visibles dans le SCI Expanded : premièrement, les villes ayant produit 75 % de la production du SCI Expanded aux deux dates et deuxièmement, les 100 villes les plus publiantes aux deux dates (2000 et 2007, en moyenne mobile sur 3 ans). Alors que l'indice de répartition est stable pour le premier sous-ensemble, il diminue nettement pour le second. Pourtant, dans le second cas, le nombre de villes est fixé à 100, ce qui veut dire que la diminution de l'indice ne peut pas s'expliquer par l'augmentation de la taille du semis. Il faut donc que les 100 villes les plus publiantes diffèrent aux deux dates. Dès lors, la baisse de l'indice suggère que les 100 villes les plus publiantes en 2007 ont tendance à être plus proches de leur plus proche voisin que les 100 villes les plus publiantes en 2000. Si cette évolution peut intéresser, elle reste très abstraite en l'absence d'une représentation cartographique de l'information.

Analyse du plus proche voisin	Villes réalisant 75% de la production (SCI Exp)		Les 100 villes les plus publiantes (SCI Exp)	
	2000*	2007*	2000*	2007*
Distance moyenne observée	236582,2	239127,9	341059,1	319749,4
Distance moyenne attendue	526090,8	534431,1	802374,3	802374,3
Index du plus proche voisin	0,45	0,45	0,43	0,40
Nombre de villes	247	282	100	100
Score Z	-16,5	-17,8	-11,0	-11,5
Part dans la production	75%	75%	54%	50%

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: SCIExp (articles, recensions et lettres)

Tableau 15 – Analyse du plus proche voisin parmi les villes les plus publiantes

Pour comparer deux semis de points, une autre méthode classique consiste à comparer la position des points de coordonnées moyennes. On obtient les coordonnées moyennes d'un semis de points sous QGIS en appliquant la commande suivante à une couche de points au format *shapefile* (« .shp ») :

Vecteur/Outils d'analyse/Coordonnée(s) moyenne(s)

En observant ce point moyen se déplacer dans le temps, on peut déterminer dans quelle direction le phénomène étudié évolue. Appliqué au semis de villes ayant publié 75 % de la production scientifique en 2000, puis à celui obéissant au même critère en 2007 (moyenne mobile sur 3 ans), on constate que le centre de gravité des villes les plus publiantes en 2007 (en jaune) est positionné plus à l'Est et légèrement plus au Sud que le centre de gravité des villes les plus publiantes en 2000 (en rouge). Ce phénomène s'explique par la « visibilisation » d'un nombre important de villes asiatiques entre les deux périodes.

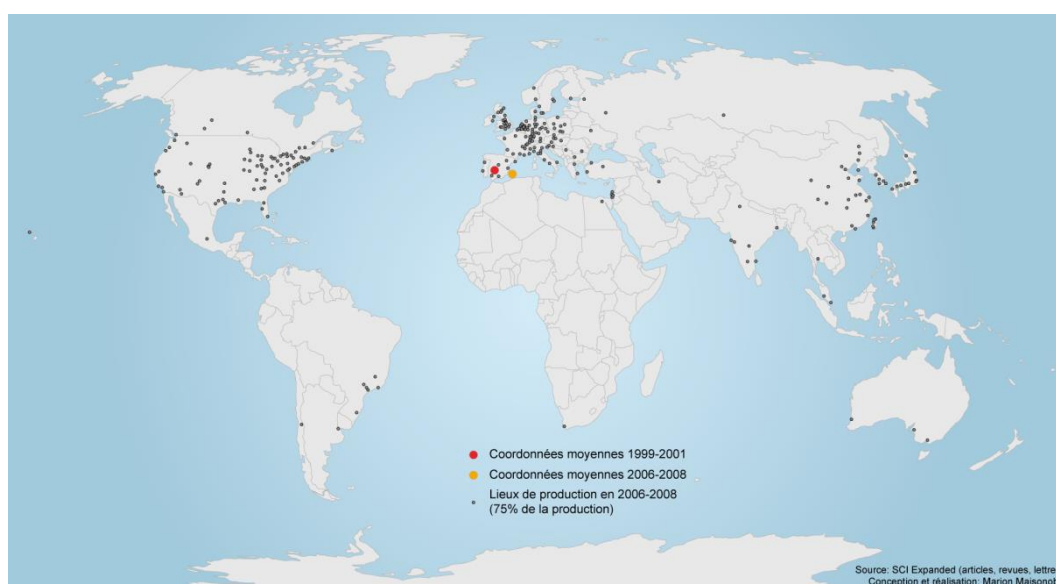


Figure 52 – Déplacement du centre de gravité de la production scientifique mondiale. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Dans les deux cas, l'analyse fait comme si l'espace considéré était continu. Or, la Figure 52 montre bien que le semis est composé de plusieurs zones très denses (en Amérique du Nord, en Europe et en Asie du Sud-Est) séparées par de grandes étendues de vide. Quelques espaces de publication excentrés se distinguent : en Amérique du Sud, en Afrique du Sud et en Australie. Cette description rapide témoigne du fait que dès qu'il s'agit de décomposer le territoire, on a recours à des découpages qualitatifs de l'espace mondial par continent et par pays. Après le recours à l'espace comme référentiel, cela revient à considérer l'espace comme support. Il s'agit de la deuxième façon de procéder pour mettre l'espace en variable. Elle s'appuie sur des connaissances implicites quant à l'appartenance des villes à des sous-espaces géographiques. Par exemple, en considérant la répartition par pays des 100 villes les plus publiantes en 2000 et en 2007 (Tableau 16), on comprend pourquoi la distance moyenne au plus proche voisin a diminué bien plus qu'on aurait pu le prévoir entre les deux périodes (Tableau 15).

En effet, le Tableau 16 signale une augmentation du nombre de villes en Chine et au Brésil parmi les 100 villes les plus publiantes entre 2000 et 2007. Ce phénomène justifie que la moyenne des distances au plus proche voisin soit plus basse en 2007. Les moyennes sont sensibles aux valeurs extrêmes or, tant qu'il n'y avait qu'une seule ville au Brésil parmi les 100 villes les plus publiantes (São Paulo), la distance de cette ville à la ville du globe la plus proche (Buenos Aires) était parmi les plus grandes distances au plus proche voisin. À partir du moment où São Paulo a été secondée par Rio de Janeiro en 2007 (dans l'ensemble des 100 villes les publiantes), alors la distance de São Paulo à sa plus proche voisine a considérablement diminué.

Pays	100 villes les plus publiantes		
	en 2000*	en 2007*	
Etats-Unis	26	25	
Chine	3	12	↗
Japon	7	6	
Allemagne	6	5	
Royaume-Uni	10	5	↘
Canada	4	4	
Italie	4	4	
Pays-Bas	4	4	
France	4	3	
Australie	3	3	
Russie	2	2	
Corée du Sud	2	2	
Espagne	2	2	
Belgique	2	1	
Taiwan	1	2	
Turquie	1	2	
Brésil	1	2	↗
Suède	4	1	↘
Argentine	1	1	
Autriche	1	1	
République Tchèque	1	1	
Danemark	1	1	
Egypte	1	1	
Finlande	1	1	
Grèce	1	1	
Hongrie	1	1	
Inde	1	1	
Suisse	1	1	
Israël	1	1	
Mexique	1	1	
Pologne	1	1	
Singapour	1	1	
Iran	0	1	
Total général	100	100	
Part dans la production mondiale	54%	50%	

*Comptage fractionné (WNC), moyenne mobile sur 3 ans.

Source: SCIEXP (articles, recensions et lettres)

Tableau 16 – Répartition des 100 villes les plus publiantes en 2000 et 2007 par pays.
Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

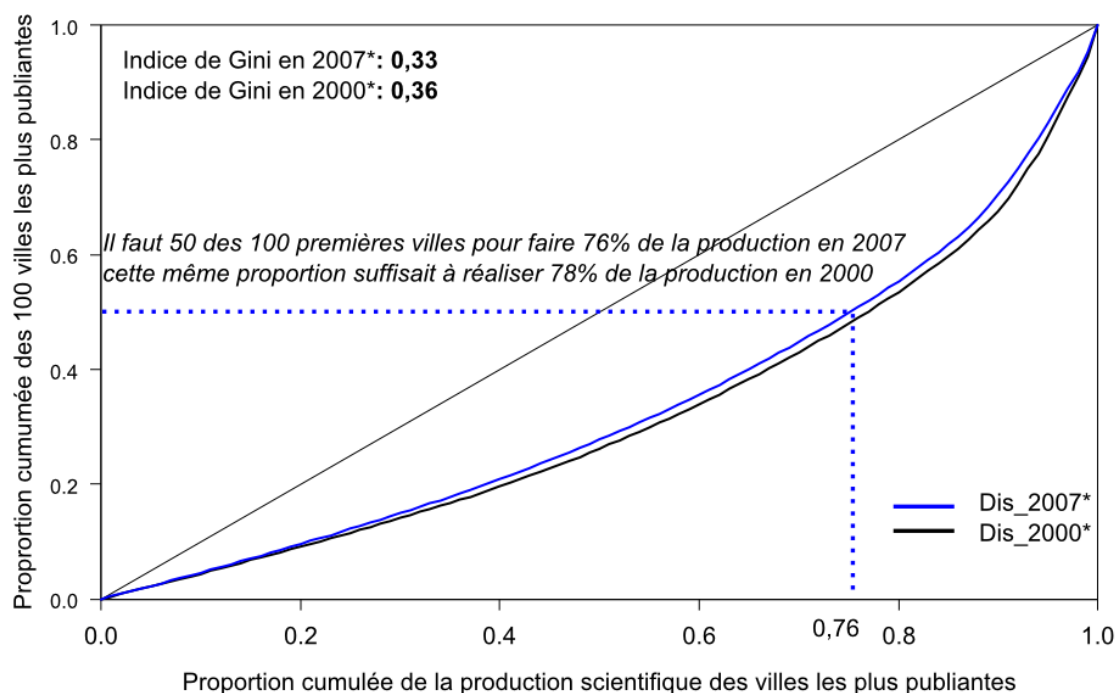
Mais la méthode du plus proche voisin interprète cette évolution comme une moindre dispersion des villes, alors que cette évolution traduit en réalité une « déconcentration » ou un « décentrage » puisque des villes éloignées des espaces denses d'Europe et d'Amérique du Nord ont fait leur apparition parmi les villes les plus publiantes, réduisant mécaniquement l'isolement des villes du « Sud » dans cet ensemble. Comme la méthode s'appuie sur des valeurs moyennes et raisonne sur une surface continue, elle est incapable de saisir ce phénomène. En fait, il faudrait l'appliquer à des sous-ensembles d'espaces pour voir des effets réels de concentration ou déconcentration.

Pour l'instant, même en se concentrant sur les villes les plus publiantes, nous n'avons pas tenu encore compte de la quantité de publications produites par les ensembles de villes étudiés. Or, la part relative de publications issue des 100 villes les plus publiantes a diminué entre 2000 et 2007. En 2000, il suffisait de 100 villes pour produire 54 % de la production mondiale alors qu'en 2007, les 100 premières villes n'arrivent plus à produire que 50 % de la production mondiale. Cela signifie que la force de production s'est répartie

un peu plus équitablement en 2007 — même si une part considérable de la production reste concentrée dans une faible quantité de villes par rapport au nombre total de villes participant à l'activité de publication (9321 villes publiantes en 2007).

Pour aller plus loin dans l'étude de ce phénomène on peut avoir recours aux indices de concentration statistique évoqués dans l'Encadré 3. Cela revient à traiter l'espace non plus comme un référentiel mais comme un support, parce que ces indices ignorent la position spatiale absolue (les coordonnées géographiques) des observations. Le plus connu de ces indices, dont nous avons déjà parlé, est l'indice de Gini, généralement utilisé pour mesurer l'inégale distribution des richesses au sein d'une population (Bellù & Liberati, 2006). Appliqué à l'ensemble des 100 villes les plus publiantes en 2000 puis en 2007, on constate que la production scientifique des 100 premières villes se répartit de façon légèrement plus équitable en 2007 qu'en 2000. En effet, alors qu'il fallait 50 villes pour effectuer 78% de cette production en 2000, les 50 premières villes ne réalisent plus que 76% de la production réalisée par les 100 premières villes en 2007, ce qui se traduit par une diminution de la valeur de l'indice de Gini (il passe de 0,36 en 2000 à 0,33 en 2007).

Courbe de Lorenz



* Moyenne mobile sur 3 ans. Source : SCI Expanded (articles, revues, lettres)

Figure 53 — La distribution de la production scientifique des 100 premières villes en 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

La courbe de Lorenz permettant de rendre compte graphiquement de ce phénomène ainsi que la mesure de l'indice de Gini ont été réalisées à l'aide du *package ineq* du logiciel R. Le logiciel R dispose de fonctions qui, appliquées à un tableau d'information géographique, facilitent l'exécution de nombreuses opérations statistiques. Par exemple, il existe des fonctions permettant, à partir d'une distribution statistique, de calculer directement un grand nombre d'indices de concentration. Ces indices sont très nombreux mais ne tiennent pas compte de la position géographique des entités spatiales. Appliqué à l'exemple qui nous occupe, cela revient à comparer la distribution des publications par entités spatiales à une distribution idéale (chaque entité contribue à part égale à la production scientifique). *Ineq* sous R permet d'obtenir la valeur de nombreux coefficients pour toute distribution statistique : Gini, Herfindahl, Rosenbluth, Theil, Atkinson, Kolm, Ricci-Schutz (également appelé Pietra) ou encore le coefficient de variation. Dans le tableau suivant (Tableau 17), nous avons reporté le taux d'évolution de plusieurs de ces coefficients par pays, c'est-à-dire en considérant la répartition géographique des publications entre agglomérations par pays. Plus précisément, on y trouve les valeurs des 4 indices de concentration les plus classiques et leur évolution calculée à partir de la répartition des publications scientifiques de l'ensemble des villes publiantes dans le SCI Expanded (articles, recensions et lettres) en 2000 et 2007 (moyenne mobile sur 3 ans).

Ainsi, il apparaît sur le Tableau 17 que le sens de l'évolution peut varier selon l'indice de concentration mobilisé. L'évolution de la valeur des indices va dans le même sens uniquement pour 7 des 20 pays considérés (les 20 pays ayant le plus publié dans le SCI Expanded en 2007). On en déduit que la production scientifique chinoise, russe, néerlandaise, turque et suédoise ainsi que la production mondiale s'est indéniablement déconcentrée entre 2000 et 2007 (les villes publiantes de ces pays contribuent de plus en plus équitablement à la production nationale) tandis que la production scientifique du Japon et de la Suisse s'est concentrée entre 2000 et 2007. L'interprétation de ces résultats occupera une partie du chapitre suivant (Chapitre 9).

En fait, sans même aller jusqu'à mesurer la valeur de ces indices, il est possible de déterminer si les publications se sont concentrées ou déconcentrées dans l'espace rien qu'en se penchant sur l'évolution de la part de publications produites par chaque agglomération, autrement dit en regardant l'évolution du poids relatif des entités spatiales dans le total de la production mondiale, par pays ou par grande région. Cette approche de l'espace comme support sera privilégiée dans les analyses qui suivront (Chapitre 9) et combinée avec une approche de l'espace en relatif à travers les liens de collaborations entre entités spatiales. Pour ces dernières analyses, de nombreuses fonctions sont également disponibles dans le logiciel R.

Indices de concentration	Taux d'évolution entre 2000* et 2007* ((x 2007-x 2000)/x 2000)				Les indices évoluent dans le même sens	Nombre de villes publiantes		Nombre de publications en 2007*
	Coeff. de variation	Herfindhal	Theil	Gini		En 2000*	En 2007*	
Etats-Unis	-3,4	-6,5	1,1	-0,2		1964	2133	272701,9
Chine	-17,4	-30,8	-2,0	-1,9	↘	225	297	90168,3
Japon	2,6	5,1	5,6	0,2	↗	207	230	74428,8
Allemagne	-0,6	-1,2	6,0	-0,1		898	1117	62172,9
Royaume-Uni	-2,7	-5,1	3,3	0,0		255	280	62139,7
France	-5,1	-9,9	3,6	-0,1		679	832	48150,3
Italie	-3,6	-6,6	4,6	-0,7		190	226	39525,9
Canada	-0,5	-1,0	2,7	0,1		198	212	37447,5
Inde	-4,1	-7,6	9,6	-1,1		340	456	30487,9
Espagne	-6,8	-12,3	3,8	-2,0		96	117	29505,0
Corée du Sud	2,4	4,6	13,5	-0,5		50	70	28621,5
Australie	1,1	2,1	16,8	-0,1		76	110	23758,1
Russie	-4,4	-8,4	-1,6	-0,9	↘	147	157	21833,6
Brésil	-10,0	-18,2	3,5	-2,9		153	218	21231,5
Taiwan	-11,0	-18,0	14,7	-3,2		16	23	19827,5
Pays-Bas	-1,6	-3,0	-1,6	-0,1	↘	98	96	18725,0
Turquie	-25,0	-42,1	-6,4	-6,2	↘	82	128	16726,9
Pologne	-6,7	-11,7	10,2	-1,1		61	79	14512,5
Suisse	2,6	5,1	5,5	0,1	↗	89	103	12967,9
Suède	-3,2	-6,0	-0,6	-0,4	↘	123	136	12642,6
Monde	-0,1	-0,2	0,0	-0,0	↘	7038	9321	1117566,3

*Comptage fractionné (WNC), moyenne mobile sur 3 ans.

Source: SCIExp (articles, recensions et lettres)

Tableau 17 – Évolution par pays des indices de concentration de la production scientifique entre 2000 et 2007. Source : SCIE Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Comme il permet à la fois de réaliser des opérations d'analyse spatiale, de statistique, et d'analyse de réseau, le logiciel R est adapté à différentes approches de l'espace géographique. Alors que le logiciel QGIS a été conçu spécifiquement pour le traitement et l'interrogation de l'information géographique, le logiciel R est, en exagérant un peu, omnipotent. R est à la fois « un langage interprété, un système d'exploration interactif et un environnement de programmation »⁸⁰. Pour réaliser des analyses quantitatives, le logiciel R propose une large gamme de possibilités. C'est un logiciel *open source*, libre et gratuit alimenté en continu par une « communauté virtuelle » de scientifiques ou d'amateurs. Son intérêt pour les chercheurs en sciences sociales n'a fait que s'accroître au cours des der-

⁸⁰ Cf. : Plutniak, S., & Maisonobe, M. (2014, mai). Travailler avec R en sciences humaines et sociales. Outils, apprentissages, possibilités. Atelier méthodologique RestTo présenté à Atelier méthodologique RestTo, Toulouse. Consulté à l'adresse <http://groupefmr.hypotheses.org/3464>

nières années. Son utilisation est facilitée par le chargement de « *packages* », c'est-à-dire d'ensembles de fonctions préalablement définies par des utilisateurs à l'aide du langage qui porte le même nom que le logiciel : le langage R. En juillet 2015, il existait plus de 6800 accessibles en ligne. La flexibilité qui en résulte fait de ce logiciel un dispositif capable d'exécuter une grande partie des traitements de données qu'un chercheur est en droit d'imaginer. Comme R offre une bonne puissance de calcul numérique mais aussi des moyens de visualisation, certains chercheurs et *a fortiori* certains géographes tendent désormais à le mobiliser pour l'ensemble de leurs analyses.

Pour aider les géographes à se saisir de ce logiciel, le manuel « R pour les géographes » a été mis en ligne en 2012 et rendu disponible en version papier depuis septembre 2014 (Groupe ElementR, 2014). Il est le fruit des réflexions et des échanges d'un groupe de huit chercheurs baptisé pour l'occasion « ElementR » et coordonné par Hadrien Commenges dans le cadre du laboratoire Géographie-cité à Paris. Le manuel passe en revue tous les besoins qu'un géographe peut avoir à combler en matière d'analyse spatiale : analyse univariée, analyses bivariées, analyses factorielles, méthodes de classification, analyse de réseaux, cartographie, statistiques spatiales. Utiliser R pour répondre à toutes ces attentes est une option séduisante mais suppose pour être retenue une très bonne maîtrise du langage. Sans cette maîtrise, l'usage de R peut avoir des effets contre-productifs. En particulier, il faut se garder d'avoir une confiance absolue dans les différents *packages* de fonctions librement implémentés. Puisqu'ils sont le fruit d'initiatives multiples, les *packages* R ne se valent pas tous et ne sont pas toujours compatibles entre eux. Par exemple, pour faire de l'analyse de réseaux sous R, il est déconseillé d'utiliser le *package statnet* en même temps que le *package igraph*, car il existe des risques de conflits entre certaines fonctions (Beauguitte, 2012). Enfin, le logiciel ne propose pas de solution adaptée à tous les besoins. En particulier, même s'il dispose d'un module de cartographie, R n'est pas du tout en mesure de remplacer un Système d'Information Géographique de type QGIS ou PostGIS (logiciels *open source*)⁸¹.

En fait, l'intérêt de R en matière de traitement de données réside dans la possibilité de reproduire, en un seul clic, une série d'opérations préalablement consignées dans un script. Mais ce mode de fonctionnement n'est pas exclusif à R, qui a d'abord été conçu pour l'exécution de traitements statistiques (ce qui en fait le concurrent historique du logiciel SAS). Pour nous, au stade de l'exploration, de l'appropriation de différentes méthodes d'analyse et de la mise au point d'un protocole, l'usage de R est complémentaire à

⁸¹ Pour preuve, la discussion qui fait suite à un billet du carnet de recherche Quanti traitant de la cartographie sous R. Le 10 Juin 2013, en réponse à « un néophyte » souhaitant réaliser des traitements géomatiques sous R, les auteurs du billet expliquait que R n'était pas « l'outil le plus adapté pour ce type de requêtes ». Cf. : Beauguitte, L., & Giraud, T. (2012, décembre 11). Cartographier avec le logiciel R [Carnet de recherche Quanti]. Consulté à l'adresse <http://quanti.hypotheses.org/795/>

celui de logiciels dont l'environnement graphique et l'architecture ont été conçus pour l'exécution de tâches plus spécialisées. En effet, les SIG comme QGIS, les logiciels de dessin vectoriel comme Inkscape ou les logiciels d'analyse de réseaux comme Pajek, Cytoscape ou Gephi ont des interfaces graphiques prévues pour s'adapter le mieux possible aux besoins de leurs utilisateurs. Loin de freiner le développement d'autres logiciels d'analyse, le développement de R se fait dans un climat propice à la cohabitation puisqu'il existe des passerelles facilitant l'échange direct de données entre logiciels. On appelle cette qualité : l'« interopérabilité ». Cette dernière permet, grâce à des codes Python, d'exécuter certaines commandes implémentées dans R depuis QGIS (Lhomme, 2012), ou bien encore d'exploiter des fonctions R depuis le logiciel Netlogo dédié à la modélisation à base d'agent⁸².

Après ces mises au point techniques, il reste encore à s'arrêter sur une catégorie particulière d'analyse largement mobilisée dans la suite de ce travail pour étudier le réseau interurbain et international des collaborations scientifiques : l'analyse de réseaux. Les outils de l'analyse de réseaux permettent d'aller au-delà d'une simple étude de la répartition spatiale de l'information, car ils offrent la possibilité de déduire une structure à partir de la position des entités spatiales dans un espace relationnel. Ce sont des outils adaptés à une conception de l'espace géographique comme support relatif. Les entités spatiales peuvent s'y mouvoir, se rapprocher ou s'éloigner les unes des autres d'après leur ressemblance ou proximité relationnelle. Il s'agit de la troisième et dernière façon de mettre l'espace géographique en variable au sens de Joël Charre.

2.2. L'organisation spatiale des activités scientifiques

En plus d'interroger la répartition géographique des publications ou des co-publications, il est possible d'aller plus loin dans l'analyse de ces dernières en mobilisant les outils de l'analyse de réseaux, dont ceux de la théorie des graphes. À partir des matrices d'adjacence construites à partir d'un ensemble de co-publications, nous pouvons extraire et analyser des graphes. Ces analyses ont pour but de nous informer sur l'organisation spatiale des activités scientifiques.

⁸² Pour faire de l'analyse de réseau sous QGIS, voir la contribution de Serge Lhomme sur le carnet de recherche du groupe fmnr : Lhomme, S. (2014, mars 11). Analyse de graphe et QGIS 2.0. Consulté à l'adresse <http://groupefmnr.hypotheses.org/3166>. Et pour une introduction plus générale au développement sous QGIS (Lhomme, 2012), voir son cours disponible en ligne : <http://serge.lhomme.pagesperso-orange.fr/doc/dev-qgis.pdf>.

Pour coupler les fonctions de R et Netlogo, voir le carnet de recherche d'Etienne Delay : Delay, E. (2014, janvier 18). Netlogo et R. Consulté à l'adresse <http://elcep.legtux.org/?p=59>

Un graphe $G = (V, E)$ est un ensemble fini et non vide de sommets (ou nœuds) V (pour *vertices*) et un ensemble fini, mais éventuellement vide, de liens (ou arêtes) E (pour *edges*) :

$$G = (V, E)$$

Un graphe se définit par son ordre (le nombre de sommets) et par sa taille (le nombre de liens). Un graphe n'est pas forcément connexe (s'il est connexe c'est que pour tous sommets, il existe au moins un chemin pour y accéder depuis n'importe quel autre sommet). Un sous-graphe connexe est appelé une composante.

Pour analyser un réseau ou graphe, trois grands types d'opérations existent :

- étudier les propriétés du graphe dans son ensemble (mesures globales) ;
- étudier les caractéristiques des sommets et, plus rarement, des liens (mesures locales) ;
- manipuler ou chercher des partitions (sous-ensembles disjoints et non-vides) qui regroupent des sommets, et plus rarement des liens, partageant des caractéristiques communes.

Nous revenons ici sur le vocabulaire et les principes de la théorie des graphes pour démontrer son intérêt analytique en matière de géographie des collaborations scientifiques. Dans cette optique, il n'est pas question d'être exhaustif sur la terminologie, les concepts et les techniques. Du reste, le lecteur qui voudrait aller plus loin dans l'appréhension du vocabulaire et des outils de l'analyse de réseaux, peut se référer aux synthèses et au lexique du groupe *fmr* en libre consultation sur internet⁸³. Les travaux du groupe *fmr* sont destinés à s'y retrouver dans un univers où le vocabulaire varie selon les spécialités et les disciplines. Précisons que l'analyse de réseaux intègre désormais un ensemble de techniques s'appuyant sur des compétences qui vont de l'algèbre linéaire aux probabilités en passant par l'algorithmique. Pour désigner cet ensemble depuis le début des années 2000, certains parlent d'une « science des réseaux » (Börner, Sanyal, & Vespignani, 2007). Malgré la très grande diversité des méthodes disponibles dans ce domaine, nous nous concentrons sur les mesures et méthodes d'analyses adaptées à l'étude des réseaux de collaborations scientifiques.

Considérons le cas des réseaux de collaborations scientifiques entre agglomérations. Les sommets sont des villes et les liens entre villes sont des fractions de contributions mesurées en tenant compte du nombre de villes signataires par publication.

⁸³ Le dépôt des synthèses en archive ouverte : <http://halshs.archives-ouvertes.fr/FMR/>
Le carnet de recherche du groupe : <http://groupefmr.hypotheses.org/>

Pour réaliser les opérations présentées dans cette thèse, nous avons souvent utilisé le logiciel Pajek. Pajek est un logiciel gratuit mais non libre développé par des mathématiciens slovènes. Bien que ce logiciel soit très complet, puissant et régulièrement mis à jour, il faut parfois, pour répondre à des besoins spécifiques, faire appel à d'autres logiciels. En particulier, certains traitements présentés plus loin ont été réalisés à partir des logiciels R (les *packages igraph* et *tnet*), Gephi, Cytoscape, VoSviewer. Les facilités de conversion des formats propres à chaque logiciel favorisent les circulations entre ces logiciels.

Pour être lues et manipulées à l'aide d'un de ces logiciels, les données relationnelles doivent être mises en forme soit dans une matrice d'adjacence (fichier que l'on charge avec une extension « .mat » dans le logiciel Pajek), soit dans une liste de liens (fichier que l'on charge avec une extension « .net » dans le logiciel Pajek). La seconde option est la moins lourde d'un point de vue informatique parce que les données redondantes n'y apparaissent pas. En effet, comme le montre la Figure 50, la matrice des collaborations scientifique est symétrique (car les liens ne sont pas orientés) et sa diagonale est vide (car les échanges intra-urbains ne comptent pas). Puisque la part des informations redondantes dans une telle matrice est importante, il est préférable de manipuler un fichier qui se présente sous la forme d'une liste de liens.

La Figure 54 montre l'allure de la liste des liens de collaborations scientifiques internationales en 2007 (moyenne mobile sur 3 ans, SCI Expanded). Chaque ligne du tableau décrit un lien entre deux pays et chaque pays est désigné par un identifiant. Dans cet exemple apparaissent l'ensemble des collaborations scientifiques internationales entretenues par le pays dont l'identifiant est 187, c'est-à-dire l'Afghanistan (les pays sont ici rangés par ordre alphabétique). En particulier, la première ligne du tableau décrit la relation scientifique de l'Afghanistan avec l'Australie (identifiant : 7). Le poids de cette relation est compris entre 0 et 1 ce qui veut dire qu'aucun article n'a été produit conjointement par ces deux pays en l'absence de la contribution d'un autre pays ou bien (les deux interprétations sont possibles en l'absence de vérification), si jamais un échange unilatéral entre les deux pays avait eu lieu entre 2006 et 2008, ce dernier n'aurait pas été suffisamment stable sur les 3 années consécutives pour que la moyenne mobile dépasse 1. Après vérification, il apparaît que la seconde interprétation est la bonne : 2 collaborations scientifiques ont eu lieu entre l'Afghanistan et l'Australie en 2006, mais elles n'ont plus été suivies par aucune collaboration en 2007 et 2008 ; de sorte qu'en lissant la valeur sur les 3 années, l'intensité de la relation scientifique entre les deux pays s'élève seulement à 0,6 (2/3).

id.source	id.destination	val.lien
187	7	0.666666666666667
187	29	0.111111111111111
187	48	0.333333333333333
187	59	0.411111111111113
187	73	0.611111111111111
187	82	0.333333333333333
187	115	0.111111111111111
187	124	0.166666666666667
187	137	0.111111111111111
187	162	0.111111111111111
187	163	0.111111111111111
187	166	0.033333333333333
187	167	0.111111111111111
187	172	0.111111111111111
187	174	0.033333333333333
187	177	1.73333333333333
187	178	3.65555555555555
1	1	0.999999999999997
1	7	0.005050505050505
1	8	0.291666666666667
1	9	0.000949667616334267
1	11	0.00740740740740733
1	14	0.000949667616334267
1	15	0.666666666666667

Figure 54 – Extrait de la liste des liens de collaborations scientifiques internationales en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Les logiciels spécialisés permettent de passer d'une matrice d'adjacence à une liste de liens sans difficulté et inversement. Sous Pajek, cette opération revient à convertir un fichier « .mat » en fichier « .net ». La Figure 55 montre comment ces fichiers doivent être composés pour être lus par ce logiciel. On observe pour le même graphe deux moyens différents de composer le fichier source. Une spécificité du logiciel Pajek ressort de cette figure, à savoir que les fichiers sources doivent être organisés en deux parties pour pouvoir être interprétés par le logiciel :

- une première partie intitulée « Vertices » (sommets) qui consiste à indiquer la liste des sommets du graphe, et éventuellement leur label et leurs attributs ;

- une seconde partie dans laquelle il faut reporter la liste des relations. Il convient alors de choisir entre deux options : la matrice d'adjacence (*Matrix*) ou la liste de liens (*Arcs/Edges*). Dans le second cas, il faut différencier les *Edges* (arêtes, c'est-à-dire les liens non orientés) et les *Arcs* (les liens orientés).

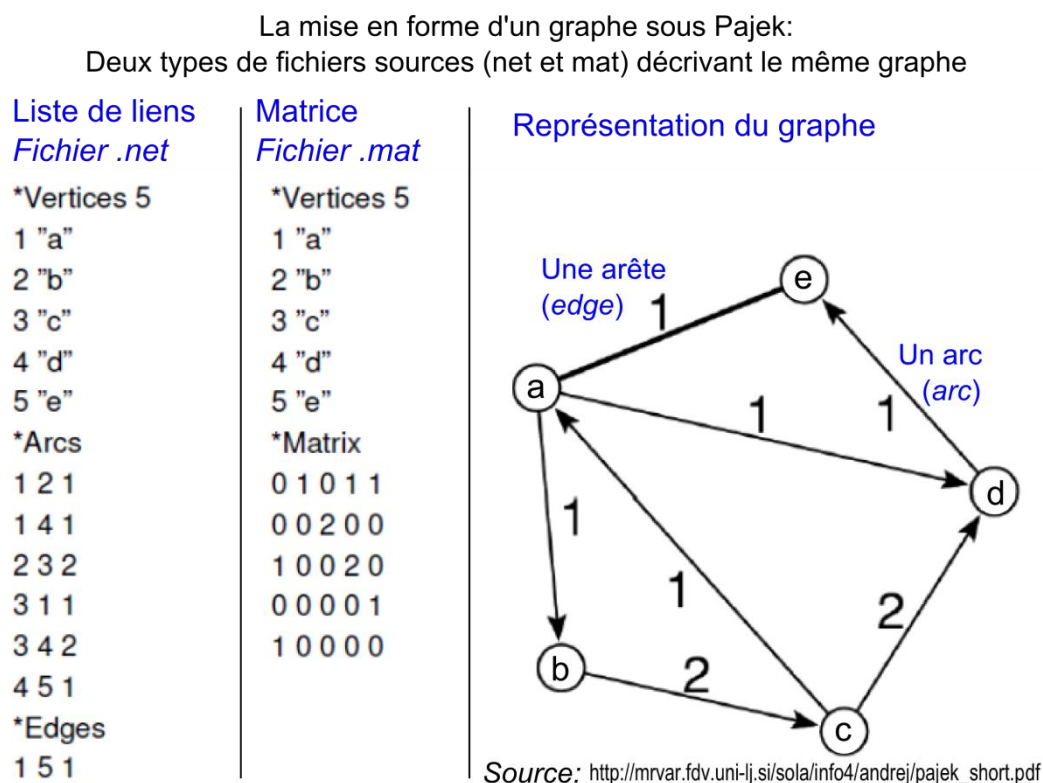


Figure 55 – La mise en forme d'un graphe sous Pajek : deux types de fichiers sources (net et mat) décrivant le même graphe. Figure reprise d'une présentation de Andrej Mrvar

Lorsqu'on charge un objet « réseau » (quelle que soit sa forme : liste de liens, matrice d'adjacence...) dans un logiciel d'analyse de réseaux, on peut accéder aux propriétés générales de l'objet considéré.

Parmi elles, nous pouvons citer :

- l'ordre et la taille du graphe ;
- le nombre de liens réflexifs (appelés *loops*) ;
- le nombre de composantes connexes et leur ordre ;
- la densité du graphe, qui est égale au nombre de liens effectifs sur le nombre de liens possibles, c'est-à-dire sur le nombre de liens qui pourraient exister dans le réseau si on avait affaire à un réseau complet ;
- le degré moyen du graphe : la moyenne des degrés de l'ensemble des sommets du graphe. Le degré est une mesure locale indiquant, pour un sommet donné, le nombre total de sommets adjacents ou voisins ;
- la distance moyenne : la longueur moyenne du plus court chemin entre deux sommets (le nombre de liens à parcourir) ;
- le diamètre : le plus long des plus courts chemins entre deux sommets.

L'ensemble de ces mesures nous renseignent sur la complétude du graphe considéré. Elles permettent de savoir si nous avons affaire à une population très connectée ou composée d'éléments faiblement reliés entre eux. Plus les distances sont longues, la densité et le degré moyen sont faibles, moins il y a de liens et plus la structure du réseau considéré est ténue.

Tandis que les mesures globales donnent une impression d'ensemble de la structure du graphe, les mesures locales renseignent la position respective de chacun des sommets. Quand certains sommets sont plus connectés que d'autres, cela suggère qu'il existe une hiérarchie entre les sommets. Parmi les mesures locales permettant d'en rendre compte, on trouve les indices de centralité dont :

- le degré d'incidence : le nombre de liens adjacents à chaque sommet ;
- la centralité d'intermédiarité ou *betweenness centrality* qui mesure le nombre de plus courts chemins passant par un sommet donné sur l'ensemble des plus courts chemins qui traversent le réseau. Elle nous renseigne sur la qualité d'entremetteur des sommets ;
- la centralité de proximité ou *closeness centrality* d'un sommet qui correspond à la somme des longueurs des plus courts chemins qui vont vers ce sommet. Elle nous renseigne sur l'accessibilité de chaque sommet. Il s'agit de l'inverse de l'indice de Shimbel (Ducruet, 2010).

Comme ces définitions le suggèrent, la plupart des indices basiques en analyse de réseaux ne portent que sur la présence ou l'absence de liens et ne tiennent pas compte de la pondération affectée aux liens, qui est caractéristique des réseaux valués. Cela revient à traiter les réseaux valués comme s'ils étaient booléens, c'est-à-dire comme si tous les liens existants prenaient pour valeur 1. Cette approche n'est pas inintéressante parce que le simple fait qu'il y ait eu au moins une co-signature entre deux villes n'est pas anodin. Toutefois, puisque nous avons privilégié un comptage entier fractionné des co-signatures, les liens de collaborations que nous considérons sont affectés d'une valeur numérique qui dépend – pour un comptage entier fractionné – de la fréquence des collaborations (le nombre total de co-publications) et de leur modalité (le nombre total de partenaires impliqués dans chaque co-publication). Cette variable quantitative peut se comprendre comme l'intensité relationnelle entre entités spatiales. Or, le différentiel entre la valeur des liens de collaborations scientifiques (dont la distribution suit une loi de Lotka) est tel qu'il est préférable d'en tenir compte pour saisir la dynamique du système étudié. Pour analyser un graphe valué, deux options, qui peuvent être combinées, sont envisageables : mobiliser des mesures pondérées qui tiennent compte de la valeur des liens (leur intensité ou force) ; et/ou appliquer un seuil au réseau pour ne traiter par exemple que les liens les plus significatifs, autrement dit, les liens supérieurs à une certaine valeur qu'il faut déterminer statistiquement, ce qui revient à ne s'intéresser qu'aux liens les plus « forts ».

Les mesures locales pondérées

Les mesures locales pondérées par la valeur des liens (première option) sont peu courantes, et présentent le défaut d'être peu sensibles à la géométrie du réseau, ou topologie. Sous Pajek, la centralité de degré est la seule mesure de centralité classique disponible dans sa version pondérée (*weighted degree*). Le degré pondéré d'un sommet correspond à la somme des valeurs des liens adjacents à ce sommet, ce qui revient à mesurer le nombre total de collaborations scientifiques par villes (le total par ligne dans une matrice d'adjacence). Conscient de cette limite en analyse de réseaux valués (mais aussi bipartites et temporels), Tore Opsahl a proposé dans sa thèse des solutions nouvelles qui ont été implémentées dans un *package* du logiciel R *tnet*, créé pour l'occasion (Opsahl, 2009). Ce *package* fonctionne en association avec un *package* plus généraliste d'analyse de réseaux appelé *igraph*, qui a intégré, depuis 2012, un certain nombre des options développées par Opsahl. Les fonctions de *tnet* (sous forme d'options dans *igraph*) permettent de calculer des indices inexistant dans Pajek. Ceux qui nous intéressent sont les indices de centralité qui tiennent compte de la valeur des liens, telle que les variantes de la centralité de proximité et d'intermédiarité appelées *weighted betweenness* et *weighted closeness*. Le problème avec ces indices (comme avec la centralité de degré pondérée), c'est qu'en intégrant davantage d'information, on prend le risque d'en oublier la géométrie du graphe.

Effectivement, il est possible pour un sommet d'avoir la même centralité pondérée tout en étant dans des positions très différentes dans le réseau. Il suffit pour lui d'avoir un seul voisin et d'entretenir une relation très intense avec lui pour être affecté d'un degré de centralité pondéré aussi élevé que celui d'un sommet doté d'un nombre très élevé de voisins, auquel il est faiblement connecté. Il existe, toutefois, quelques indices de centralité qui permettent d'échapper à cette limite : la *centrality combined*, l'*eigenvector centrality* et la *laplacian centrality*. Ces mesures sont encore très rarement utilisées, quand bien même elles sont plus sensibles que les autres à la place des sommets dans le graphe.

La première, la centralité combinée, proposée par Tore Opsahl, met en rapport le degré pondéré et le degré d'incidence selon les termes attribués à un coefficient α . Pour reprendre l'explication de César Ducruet : « le réglage d' α (entre 0 et 1) dépend de la thématique abordée : un coefficient faible donnera plus d'importance au degré d'incidence, tandis qu'un coefficient fort favorisera le rôle du poids de liens » (Ducruet, 2010). Les deux dernières mesures reposent sur des méthodes d'algèbre linéaire : la recherche de valeurs propres dans une matrice ou *eigenvector* pour la première et le calcul de la matrice laplacienne pour la seconde.

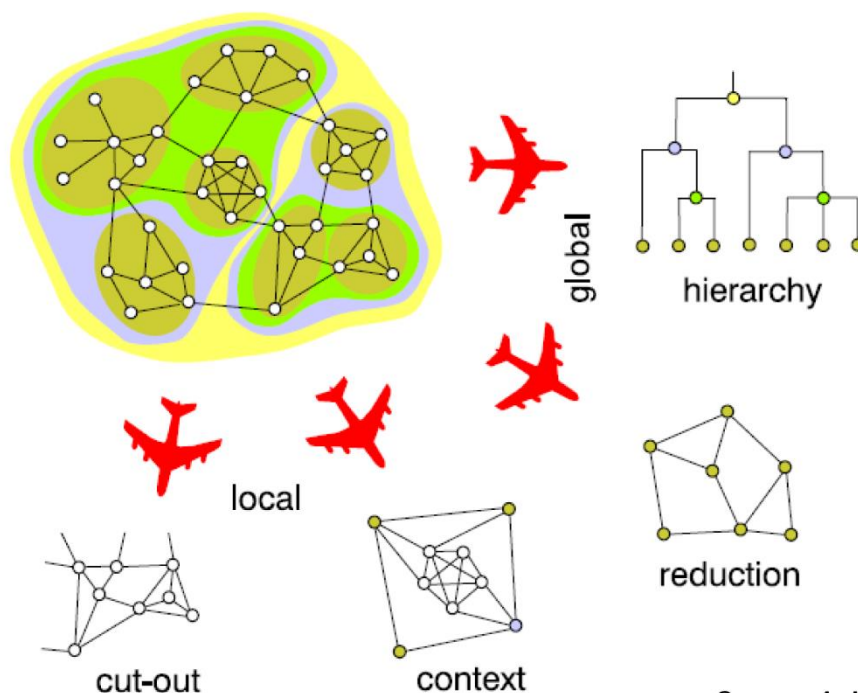
Dans sa thèse, pour étudier des partenariats en recherche et innovation dans le cadre de pôles de compétitivité, Grandclément a privilégié l'*eigenvector centrality* parce

qu'elle tient compte de la place de chaque nœud dans l'ensemble du graphe et qu'elle valorise la centralité d'un nœud, non pas du point de vue de son intermédiarité (le nombre de chaînes passant par ce nœud) mais du point de vue de son accessibilité (Grandclément, 2012, p. 312). Plus précisément, en s'appuyant sur les travaux de Bonacich, Freeman, Hanneman et Riddle ainsi que Newman, Grandclément a expliqué qu'« au lieu de ne compter que les liens formés par un acteur, [cette mesure] introduit une distinction entre les liens, en considérant qu'une arête vers un nœud central a plus de valeur qu'une arête vers un nœud périphérique. Ainsi, la centralité d'un nœud dépend à la fois du nombre de ses connexions et de la qualité de ses connexions, c'est-à-dire de la centralité des nœuds auxquels il est relié. Avoir de nombreuses relations reste important mais un nœud possédant un petit nombre de relations-clés sera aussi considéré comme central. » (Grandclément, 2012, p. 313).

La dernière mesure, la centralité laplacienne (*laplacian centrality*) a été définie encore plus récemment et n'a pratiquement pas été utilisée depuis (à notre connaissance). Elle se fonde sur la matrice laplacienne (la différence entre la matrice des degrés et la matrice d'adjacence) et mesure la centralité de chaque nœud d'après l'influence que leur suppression aurait sur la position des autres nœuds et plus précisément sur l'optimisation d'un indicateur d'« énergie laplacienne » calculé à partir de la matrice laplacienne (Qi, Fuller, Wu, Wu, & Zhang, 2012 ; Qi *et al.*, 2013). Si cette dernière mesure a été implémentée dans le logiciel Pajek en 2014, il convient de la considérer avec réserve tout comme l'*eigenvector centrality* car des spécialistes ont pointé d'une part le fait que les mesures se fondant sur les valeurs propres sont très sensibles à des variations minimales dans la structure du graphe, et d'autre part que les matrices laplaciennes diffèrent d'autant plus entre elles que la distribution de degré est très hiérarchisée (comme c'est le cas avec les réseaux de collaborations scientifiques) (Beauguitte & Beauguitte, 2011).

Plutôt que de se fier à ces mesures locales, il existe une approche alternative et/ou complémentaire permettant d'être sensible à la géométrie ou structure du graphe (seconde option). Il s'agit de considérer des sous-graphes ou partitions. Les partitions peuvent être des sous-ensembles de sommets ou des sous-ensembles de liens. En analyse de réseaux, à partir de partitions, ou sous-ensembles disjoints d'éléments partageant des caractéristiques communes, il est possible d'effectuer un grand nombre d'opérations. C'est souvent sur des partitions de sommets que l'on se fonde pour décomposer, simplifier ou réduire un graphe. Pour se risquer à faire un parallèle avec le contenu de la section précédente (2.1), choisir de s'intéresser à la valeur des mesures locales sur l'ensemble du graphe ou préférer commencer par découper le graphe en sous-graphe, dépend de l'approche analytique que l'on privilégie.

Vues locales et globales d'un réseau



Source: A. Mrvar

Figure 56 – Vues locales et globales d'un réseau. Figure reprise d'une présentation de Andrej Mrvar portant sur le logiciel Pajek

Toutefois, chacune des approches n'exclut pas l'autre. Le fait de différencier les sommets d'un graphe offre la possibilité de faire varier la focale ou l'échelle d'analyse :

- se concentrer sur un groupe spécifique de sommets et les relations qu'ils entretiennent à l'intérieur de ce groupe (approche locale) ;
- se concentrer sur les relations entre les différents groupes. Chaque groupe de sommets peut être traité comme un sommet une fois que le graphe a été réduit (approche globale) ;
- combiner l'approche locale et globale, par exemple en regardant dans le détail comment des sommets appartenant à un groupe donné s'inscrivent dans le contexte d'ensemble que décrit le graphe (approche contextuelle).

Ces différentes options sont résumées dans la Figure 56 issue d'un diaporama réalisé par Andrej Mrvar (l'un des développeurs de Pajek), accessible depuis son site personnel⁸⁴.

⁸⁴ Source : URL : http://mrvar.fdv.uni-lj.si/sola/info4/andrej/pajek_short.pdf

Le Partitionnement d'un graphe

En matière de partitionnement (opération qui consiste généralement à associer une partition à un sommet), il existe deux façons de procéder : mobiliser une partition connue *a priori* ou bien définir une partition en s'appuyant sur la structure du graphe.

Dans le premier cas et avec l'exemple du réseau des collaborations scientifiques interurbaines, il peut s'agir d'associer son pays d'appartenance à chacune des villes. Ainsi, on se retrouve en mesure d'analyser :

- la structure relationnelle interne à chaque pays ou structure nationale (approche locale) ;
- la structure internationale, en réduisant, par exemple, le graphe de départ à celui des collaborations entre pays (approche globale) ;
- la structure transnationale, en analysant, par exemple, les liens entretenus entre les villes d'un pays donné et les autres pays du monde (approche contextuelle).

Ici, l'appartenance nationale est connue à l'avance et ne dépend pas de la place des sommets dans le graphe considéré. Il est intéressant de la mobiliser si jamais nous souhaitons vérifier des hypothèses sur le caractère structurant de cette information pour le réseau. D'ailleurs, comme la majorité des travaux de scientométrie spatiale prennent le niveau national comme niveau de résolution, nous verrons qu'il est nécessaire d'en tenir compte pour nos analyses et nos conclusions. En particulier, s'intéresser à la structuration nationale des collaborations scientifiques permet de confronter les données à un grand nombre de discours en vogue sur la « mondialisation » de la science, souvent perçue comme synonyme de l'internationalisation des activités scientifiques, on y reviendra.

L'autre possibilité en matière de partitionnement présente aussi beaucoup d'intérêt pour analyser le réseau des collaborations scientifiques. Elle consiste à rechercher des partitions qui sont déterminées par la position relative des sommets dans le graphe. Cette approche est utile pour décomposer la structure d'ensemble, l'armature du réseau ou « décor ». Une fois le « décor » mis à nu, alors il est plus facile, en analysant les variations du « trafic » (les flux ou la fréquence des relations) à plusieurs dates, de distinguer les observations qui dépendent de l'organisation courante du système et celles qui sortent de l'ordinaire, qui témoignent d'un changement, d'une évolution dans la routine.

La plupart du temps, rechercher des partitions consiste à repérer les parties du graphe qui sont très cohésives, c'est-à-dire trouver des sous-graphes denses à l'intérieur d'un graphe (des sous-graphes à l'intérieur desquels l'ensemble des connexions possibles est réalisé). En analyse des réseaux sociaux, on appelle traditionnellement ces sous-graphes, des « cliques ». Parmi ces cliques, les sociologues s'intéressent tout particulièrement aux « triades ». Pour simplifier, disons qu'il suffit d'être deux pour qu'il y ait une

relation (on parle de dyade), mais c'est généralement à partir du moment où l'on est trois que l'on entre en présence d'un groupe social ou collectif et, plus concrètement, que l'on peut se poser la question de la transitivité : les amis de mes amis sont-ils mes amis ?

Exactement comme pour les mesures locales, les méthodes de recherche de partitions les plus classiques n'ont pas été conçues pour s'appliquer aux graphes valués. Au contraire, dans le cadre des études sociométriques, l'orientation des liens et leur signe (positif ou négatif) a été très tôt pris en compte (White, Boorman, & Breiger, 1976). En effet, c'est en s'appuyant sur l'orientation des liens qu'on peut le mieux déduire des positions structurellement équivalentes de type : dominant/dominés (Chase, 1982), ou centre/périphérie (Smith & White, 1992). Avec les liens de collaborations scientifiques, cette approche est plus difficile à adopter puisque les relations ne sont pas orientées. En fait, la démarche la plus adaptée pour partitionner le réseau des collaborations scientifiques consiste à trouver des groupes de villes fortement connectées entre elles en s'appuyant sur la valeur de leurs relations. Ainsi, ce sont en priorité les techniques adaptées aux réseaux valués qui nous intéressent. Cependant, si l'on devait analyser un réseau de citations scientifiques entre villes ou un réseau de liens majeurs ou préférentiels entre villes pour en déduire des relations de dépendance, alors on pourrait recourir aux méthodes de partitionnement adaptées aux réseaux dirigés.

Pour partitionner un graphe d'après sa structure, on peut :

- décider *a priori* en combien de morceaux ou partitions diviser le graphe ;
- s'en remettre à un algorithme qui détermine le nombre de groupes dont la cohésion satisfait des conditions optimales. Pour cela, on se ramène de plus en plus souvent à un problème de probabilité (on mesure la différence entre la structure observée à celle d'un graphe aléatoire).

La détection d'objets appelés *islands* (îles) accessible sous Pajek est proche d'une méthode du premier type puisque le résultat dépend d'une contrainte sur l'ordre des partitions à trouver qu'il faut fixer *a priori*. On y accède dans le logiciel en allant à :

Network/Create Partition/Island

Après avoir fixé le nombre maximal de sommets que l'on souhaite regrouper dans la même « île », l'algorithme détecte des groupes obéissant à la contrainte suivante : la valeur du lien minimale dans l'île doit être plus grande que la plus grande valeur de lien se trouvant au voisinage de l'île. On obtient ainsi des sous-groupes connexes dont les sommets sont davantage connectés entre eux qu'avec leurs voisins (Batagelj, Kejzar, & Korenjak-Cerne, 2006). Si l'on ne fixait pas ce paramètre (l'ordre de l'île) *a priori*, alors il pourrait arriver que l'ensemble de la composante principale du réseau constitue une île. Pour filer la métaphore insulaire, la fixation du paramètre revient à fixer le niveau de l'eau sur une surface dont le relief varie en fonction de l'intensité des liens au voisinage de chaque

sommet. Une manière d'améliorer l'efficacité de cette méthode consiste à tester plusieurs ordres d'île pour retenir le plus intéressant, autrement dit, à réitérer l'opération plusieurs fois jusqu'à obtenir le découpage le plus satisfaisant (Nooy, Mrvar, & Batagelj, 2005).

Si elle est en apparence plus robuste, l'option probabiliste (la seconde option qui laisse à l'algorithme le soin de déterminer le nombre de partitions) pose bon nombre de problèmes qui tiennent à la fixation des paramètres guidant l'opération d'optimisation. Ces problèmes n'épargnent pas les méthodes les plus récentes qui ont été mises au point, souvent par des informaticiens, et qualifiées par ces derniers de méthodes de *clustering* ou « détection de communautés ». Parmi les plus populaires, la méthode de Louvain (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008 ; Blondel, Krings, & Thomas, 2010) est implémentée dans le logiciel Pajek depuis 2012. La méthode de Louvain est également disponible dans les logiciels Gephi et NetworkX ainsi que dans le package *igraph* du logiciel R. L'algorithme est plus rapide sur Pajek que sur Gephi. Dans Pajek, on la trouve à : **Network/Create Partition/Communities/Louvain method**

L'algorithme qu'elle mobilise est adapté aux graphes valués, y compris s'ils sont signés. Comme pour beaucoup d'algorithmes de détection de communautés, la qualité du partitionnement est optimisée par maximisation de la fonction de modularité (Fortunato, 2010). L'indice de modularité mesure la différence entre la densité d'un graphe (ou sous-graphe) donné et la densité d'un graphe aléatoire possédant les mêmes caractéristiques (même nombre et poids des liens). Maximiser la fonction de modularité pour partitionner un graphe revient à s'assurer que le nombre et le poids des liens est plus important à l'intérieur des partitions qu'entre les partitions. Pour l'exprimer autrement, il faut que la densité intra-communautaire dépasse la densité inter-communautaire.

La méthode de Louvain est stochastique, *bottom-up* et multi-niveau. Au lancement de l'algorithme, tous les sommets appartiennent à une partition différente. Ils sont regroupés, par itération, dans des partitions de modularité optimale. Arrivé au premier optimum, le processus se poursuit au niveau supérieur : chaque partition est traitée comme un sommet et ainsi de suite. L'opération continue jusqu'à ce qu'il n'y ait plus aucun gain de modularité possible. En exécutant l'algorithme à plusieurs niveaux, on cherche à échapper au défaut de la modularité connu sous le nom de « limite de résolution » (Fortunato & Barthélemy, 2007). Si le fait d'associer une partition du premier niveau avec une autre devait engendrer une perte de modularité, cette association ne serait pas réalisée par la méthode de Louvain. Ainsi, les petites structures dotées d'une meilleure modularité ne risquent pas d'être noyées dans des structures plus grandes et moins significatives ; contrairement à ce qui risquerait d'arriver avec un algorithme *top-down* appliqué à un trop grand graphe.

Le défaut reconnu de cette méthode provient de la sensibilité du résultat à l'ordre de traitement des sommets, car il a une influence sur le partitionnement. De ce fait, il peut être utile de procéder à des permutations pour s'assurer de la robustesse du résultat (Blondel *et al.*, 2010). L'autre piste est d'imposer un ordre qui tienne compte des caractéristiques du réseau (De Meo, Ferrara, Fiumara, & Provetti, 2011). Plutôt que l'algorithme classique de Louvain, Pajek propose une amélioration avec raffinement multi-niveau qui tend à donner de meilleurs résultats pour le partitionnement des grands graphes (Rotta & Noack, 2011).

En même temps que la méthode de Louvain a été intégrée au logiciel Pajek, une méthode voisine a été proposée qui s'appelle VoS. La méthode VoS optimise une autre fonction de qualité que la fonction de modularité : la fonction VoS *clustering* (Waltman *et al.*, 2010). Cette fonction est une variante pondérée de la modularité qui a été pensée pour le traitement des données bibliométriques. En l'appliquant à la place de la modularité, on donne moins de poids aux sommets de très haut degré. Cette méthode de détection de communautés convient aux graphes pour lesquels la valeur des liens rend compte de l'importance de la similarité entre sommets. Pour compléter leur méthode, les concepteurs, Nees Jan van Eck et Ludo Waltman, ont cherché quel serait le meilleur moyen de visualiser le graphe compte tenu du partitionnement obtenu. En d'autres mots, ils se sont occupés de mettre au point un algorithme de visualisation qui puisse positionner les sommets dans un plan en deux dimensions en tenant compte de leur similarité, et plus particulièrement en s'appuyant sur le même critère de similarité que celui utilisé pour le partitionnement. Plus les sommets se « ressemblent » (ont une relation privilégiée mais aussi entretiennent des relations privilégiées avec les mêmes sommets), plus ils devront être placés à proximité l'un de l'autre sur la représentation graphique obéissant à l'algorithme VoS. Cette approche unifiée permet d'éviter qu'il y ait des incohérences entre le résultat du processus de *clustering* et la représentation graphique de ce résultat.

Naturellement, l'interface de visualisation de Pajek (accessible depuis le menu **Draw**) propose aussi l'algorithme de visualisation VoS. Il fonctionne sur le principe de l'optimisation de la fonction de qualité VoS *mapping*. Étant donné la proximité des deux méthodes de partitionnement, le VoS *mapping* s'adapte aussi à la représentation d'un graphe partitionné suivant la méthode de Louvain. Une comparaison des deux méthodes de classification est accessible sur le site personnel d'Andrej Mrvar⁸⁵. Des précisions sur leur utilisation s'y trouvent également, notamment en ce qui concerne : le paramétrage, l'évaluation de la qualité du partitionnement, et la représentation graphique du résultat. Enfin, depuis Pajek, une passerelle existe vers le logiciel VoSviewer qui offre la possibilité d'obtenir rapidement une représentation graphique interactive et élégante du résultat.

⁸⁵ Source : URL : <http://mrvar.fdv.uni-lj.si/pajek/community/LouvainVOS.htm>

Ci-dessous, la Figure 57 donne un exemple de visuel obtenu à partir des données de collaborations scientifiques entre municipalités néerlandaises (toutes disciplines et tous types de publications confondus). La couleur des sommets dépend du groupe auquel ils ont été affectés.

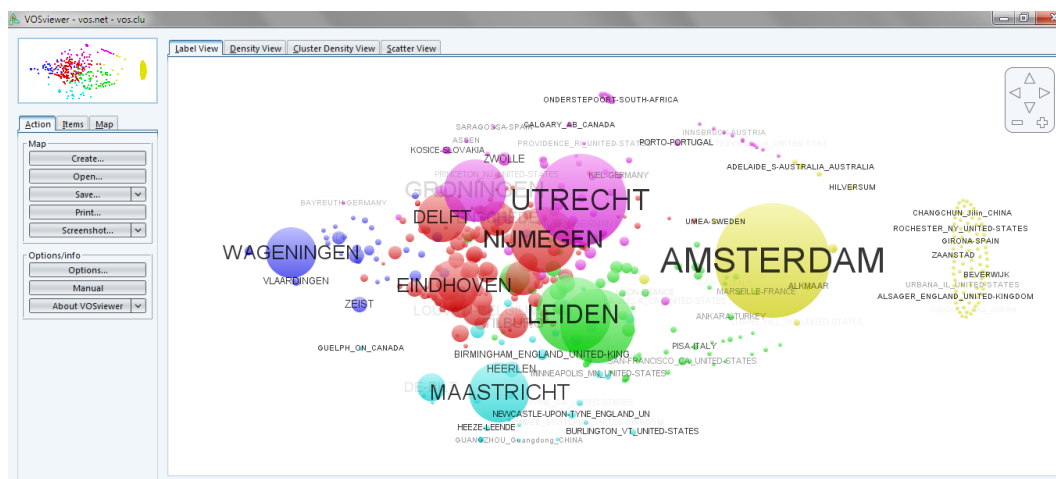


Figure 57 – L'interface de visualisation VoSviewer. L'exemple des collaborations scientifiques néerlandaises. Source : Web of Science, 2006-2008

Notons que le résultat obtenu est lisible et interprétable en s'abstenant de faire figurer les liens. En effet, puisque la position des sommets dépend de leurs relations à tous les autres sommets, c'est cette position qu'il importe de visualiser. Quand les liens sont très nombreux, leur présence complique davantage qu'elle n'améliore la lisibilité de la représentation. Il faut préciser que les développeurs n'ont pas prévu de moyens d'exporter ce rendu graphique dans un format vectoriel, qui est pourtant le format préférable pour produire une légende et réaliser les ajustements visant à mettre aux normes les sorties graphiques. C'est pour cela que nous mobilisons essentiellement ce logiciel comme outil d'exploration et de calcul.

En fait, l'algorithme de visualisation VoS s'appuie en pratique sur les techniques d'analyses multivariées et les solutions aux problèmes classiques de positionnement multidimensionnel (MDS). Ces problèmes ont été formalisés pour la première fois au milieu des années 1950 par le mathématicien, statisticien et informaticien Joseph Kruskal (dont la distance de collaboration à Paul Erdős vaut deux⁸⁶). Ces problèmes font appels à des mesures de similarité, aux méthodes d'optimisation et de classification ainsi qu'aux tech-

⁸⁶ D'après les données du projet de l'université d'Ockland autour du nombre d'Erdős : le collaborateur commun à Erdős et Kruskal est le mathématicien Allan J. Hoffman (Source : URL : <http://www.oakland.edu/enp/thedata/>). Bien que Paul Erdős n'ait pas directement collaboré avec Kruskal, l'histoire raconte qu'il a joué un rôle très important dans sa thèse.

niques de visualisation (Kruskal, 1964). Les techniques de MDS mises au point par Kruskal fonctionnent sur le principe des analyses en composante principale. À notre connaissance, les chercheurs Frame et Carpenter ont été les premiers à appliquer cette méthode sur une matrice de co-signatures entre pays (obtenue à partir des publications indexées dans le SCI) (Frame & Carpenter, 1979). Pour expliquer leur approche, ces derniers se réfèrent au livre de Kruskal et Wish paru un an plus tôt, *Multidimensional Scaling* (Kruskal & Wish, 1978). Pour le calcul, Frame et Carpenter se servent d'un programme développé par les laboratoires Bell appelé M-D SCAL. Parmi les inconvénients qu'ils soulignent à l'issue de leur analyse, il y a le fait qu'il faut indiquer le nombre de groupes ou dimensions à l'avance. Depuis lors, cette limite a été dépassée, les techniques d'analyses multivariées se sont multipliées, elles se sont adaptées à l'accroissement de la vitesse de calcul en informatique et ont été testées dans plusieurs domaines, dont la scientométrie.

Généralement, les mesures de similarité ou proximité sur lesquelles reposent ces méthodes de visualisation contrôlent par la taille (le nombre total de collaborations associé à chaque sommet du graphe) suggérant implicitement qu'il s'agit d'un facteur déterminant dont il faut se défaire pour mettre à jour les véritables affinités entre entités d'analyse. Pour cette raison, ces approches sont déjà dans le registre de l'explication des observations et non plus seulement dans celui de la description. En analyse spatiale, expliquer revient à passer de la question « Où comment ? » à la question « Où pourquoi ? ».

2.3. Expliquer la géographie des activités scientifiques

Comme l'indiquent nos résultats préliminaires, les sciences sont produites quelque part et il fait sens d'en étudier la géographie, d'autant que cette géographie évolue dans le temps. La répartition des activités scientifiques au même titre que la répartition des sabots est loin d'être d'aléatoire, et c'est sans doute ce qui la rend digne d'intérêt d'un point de vue géographique. L'idée de cette comparaison nous vient d'un passage de l'œuvre de Mark Jefferson, repris par Pierre George, qui démontre l'intérêt d'analyser des répartitions géographiques, y compris la répartition des sabots, sans en rester à de simples descriptions. Pour Jefferson : « Aucune répartition sur la terre n'est due au hasard. Aucune chose n'est indépendante d'une autre. Une certaine manière d'ordre est généralement saisie dans cette répartition, même si l'on n'en perçoit pas tous les détails » (George, 1970). Au sujet des sabots, il illustre : « Les sabots ont été originellement faits par des artisans sachant utiliser des outils en fer. Par conséquent, on ne les trouve que dans les pays ayant acquis la technique du travail du bois. Mais en même temps, leur fabrication est en rapport avec la répartition des climats pluvieux et humides générateurs de forêts. Les cartes de répartition des hommes sachant se servir d'outils en fer, des forêts et des

climats pluvieux, nous mettent en mesure de prévoir où nous allons trouver des sabots. Mais il y a des exceptions » (*ibid.*). La répartition des activités scientifiques comme celle des sabots est susceptible de s'expliquer, y compris par des facteurs de géographie physique puisque, par exemple, les sciences océanographiques sont plus volontiers pratiquées en bordure de littoral — à quelques exceptions près, qui doivent pouvoir s'expliquer, comme l'Institut Océanographique de Paris, inauguré en 1911 par le prince Albert de Monaco (Académie des Sciences, 1922). Mais alors que les sabots ont trop de points communs avec d'autres biens de consommation ordinaire pour nécessiter un cadre théorique et des méthodes d'analyse spatiale spécifiques, il nous semble que le domaine des sciences est suffisamment vaste et autonome à l'égard des autres faits humains pour qu'il soit raisonnable de réfléchir à des moyens génériques de l'étudier en géographie.

La différence tient au nombre et à la variété des objets scientifiques, à la particularité des sources permettant de les étudier, voire enfin à leur capacité à se comporter selon des logiques propres ou normées. Ainsi, la science comme objet d'étude est composée de sous-catégories telles que les disciplines, les diasporas, les spécialités, les mouvements, les questions de recherche, les laboratoires, les conférences etc., dont on peut étudier les géographies. Compte tenu de la diversité des cas que recouvre le concept de science, il est pertinent de réfléchir à des façons d'en généraliser l'étude, mais aussi de tester des régularités. Ce paragraphe ne vise pas à discréditer la géographie des sabots, mais simplement à remobiliser la conviction, largement défendue lors de l'état de l'art (Partie 1), selon laquelle il manque à la géographie une branche pour aborder et traiter les objets particuliers que recouvre la « science ». Quand cette branche sera consolidée alors de nombreuses voies seront ouvertes comme, par exemple, le développement des études géographiques de spécialités de recherche. L'étude géographique de la spécialité de la réparation de l'ADN, qui occupe la dernière partie de la thèse, procède de cette logique (Partie 3). Faire la géographie d'une spécialité scientifique en s'appuyant sur les données de production nous a permis de tester la méthode générique mise en place dans le cadre du programme « Géoscience », mais aussi de mobiliser différents moyens comme les entretiens et les récits de chercheurs pour comprendre l'émergence et la diffusion spatiale des recherches portant sur la réparation de l'ADN. Si ce type d'études allait en se multipliant, comme l'indique la tendance depuis le milieu des années 2000, alors un domaine de recherche bientôt aussi balisé que la géographie économique pourrait être emprunté sans crainte, celui de la géographie des sciences.

Comme elle permet une compréhension plus fine des phénomènes, l'étude de cas est un bon moyen d'identifier les facteurs déterminants pour la géographie du domaine de recherche considéré. Elle se nourrit d'observations qualitatives qui enrichissent et expliquent la géographie de la production scientifique du domaine de recherche analysé et son

évolution à l'échelle mondiale. Ce faisant, elle ouvre des pistes quant aux déterminants de la localisation et de l'organisation spatiale des activités scientifiques en général. Cette démarche est néanmoins inséparable de la démarche plus quantitative sur laquelle se fonde le prochain chapitre. En effet, les résultats de l'étude de cas (Partie 3) sont précédés (Chapitre 9) d'une analyse quantitative de la portion la plus complète et la plus significative du *Web of Science* : les articles, recensions, et lettres indexés dans le SCI Expanded en 1999-2001 et 2006-2008 (soit le cumul de plus de 5 millions de publications). Ce travail, fruit de nos collaborations, permet de capturer les grandes tendances contemporaines traversant la géographie des activités scientifiques dans le monde à plusieurs échelles. La connaissance de ces grandes tendances fournit un cadre pour interpréter les observations qui s'inscrivent dans le contexte contemporain de croissance de la production scientifique à l'échelle mondiale, au rang desquelles figure l'expansion géographique du domaine de la réparation de l'ADN.

Ces grandes tendances sont déduites de la description des observations spatialisées réalisée au moyen des différents outils d'analyse quantitative que l'on vient d'évoquer : représentations graphiques, analyses de distributions statistiques, analyses de réseaux. Le déroulement, les choix opérés et les interprétations accompagnant cette démarche descriptive participeront de l'explication. Mais, comme il existe des approches statistiques qui relèvent plus directement de l'explication, nous traitons plus avant de cette question dans les derniers paragraphes de ce chapitre. En prolongeant l'exemple de la géographie quantitative des productions et des collaborations scientifiques dans le monde, la question « Où pourquoi ? », abordée ici, porte sur les causes de la répartition géographique qui peuvent être à la fois internes et externes à l'activité scientifique.

Les causes internes sont d'abord celles que l'on peut déduire des données elles-mêmes. Si l'on prend l'exemple d'une matrice de collaborations scientifiques, il est courant parmi les spécialistes des sciences de distinguer deux variables explicatives majeures :

- le poids ou la taille des entités d'analyse (par exemple, le nombre de publications ou de co-publications scientifiques par entité spatiale) ;
- la similarité ou proximité des entités d'analyse (l'intensité de leurs relations respectives indépendamment de leurs poids).

Classiquement, cette seconde variable est jugée plus intéressante à mettre en évidence que la première, qui serait pour ainsi dire purement structurelle. Entre 2005 et 2010, au moment de développer la méthode VoS, les chercheurs en *Science and Technology Studies* à l'Université Erasmus de Rotterdam, van Eck et Waltman se sont donc penchés sur la nécessité de contrôler par l'effet de taille avant d'analyser des données relationnelles.

Ils reviennent ainsi sur plusieurs décennies de recherches en scientométrie fondées sur l'analyse de matrices de co-occurrence (de mots, de citations...) et plus rarement de co-signatures (d'auteurs, d'institutions, de pays...) pour rediscuter des deux questions classiques dans le domaine de la recherche d'information bibliographique (ou *information retrieval*) :

- Pourquoi faut-il normaliser les données de co-occurrence ? (Waltman & Eck, 2007)

- Quelles mesures de similarité faut-il mobiliser pour le faire ? (van Eck & Waltman, 2009)

D'après eux, le nombre de co-occurrences de deux objets peut être vu comme le résultat de deux effets indépendants : l'effet de taille et l'effet de similarité. Pour van Eck et Waltman, l'opération de normalisation doit permettre d'annuler entièrement l'effet de taille. L'idée est d'établir un degré de similarité entre les objets quelle que soit leur taille, autrement dit, comme si tous les objets avaient la même fréquence d'apparition ou probabilité de se manifester. C'est ce qui explique que parmi les indices directs ou locaux, ils jugent que les mesures probabilistes sont les mieux adaptées. En effet, plusieurs des indices les plus populaires (Jaccard, Salton, Cosinus) ne permettent pas, d'après eux, d'éliminer entièrement l'effet de taille. Ces indices sont fondés sur l'idée de mesurer le recouvrement relatif entre deux objets. Dans ce cas, la valeur de l'indice pour deux objets donnés (ou entités spatiales) i et j , dépend de la part de co-occurrences (ou co-signatures) de i avec j sur le total de leurs occurrences respectives. Ces indices ne tiennent pas compte de l'écart entre le nombre de co-occurrences observées et le nombre qu'on obtiendrait dans une situation d'indépendance statistique, c'est-à-dire au cas où tous les objets seraient de même taille. Van Eck et Waltman estiment, pour leur part, que ce n'est qu'en référence au modèle d'indépendance statistique qu'il est possible de raisonner toutes choses égales par ailleurs. Ils préconisent de ce fait l'utilisation de l'indice dit de proximité, de force ou d'affinité (l'appellation diffère selon les auteurs) qui obéit à la définition suivante pour s_i le total des co-occurrences de l'entité i , s_j le total des co-occurrences de l'entité j , et S_A le degré de similarité entre i et j :

$$S_A(c_{ij}, s_i, s_j) = \frac{c_{ij}}{s_i s_j}$$

Il s'agit d'un ratio entre le nombre de co-occurrences observées et le nombre de co-occurrences attendues en situation d'indépendance statistique. Ils suggèrent d'en multiplier le résultat par le nombre de publications total (la somme des marges de la matrice d'adjacence dans le cas où le comptage des publications est fractionné).

Mais, la scientométrie n'est pas la seule discipline à se préoccuper de contrôler par l'effet de taille avant d'analyser le contenu d'une matrice. De la même manière, en analyse spatiale, il est important de contrôler par l'effet de taille pour capturer l'effet joué par la distance géographique sur les interactions ou flux (Grasland & Baron, 2006). Face à une matrice de flux, il convient d'estimer l'effet conjugué de la taille et de la distance à l'aide d'un modèle dit gravitaire ou d'interaction spatiale. L'opération permet de déduire la valeur des constantes de la loi gravitaire qui s'applique au phénomène étudié. Aussi, retrouve-t-on cette approche en science régionale où la similarité entre deux entités spatiales se traduit non seulement par leur « distance » (au sens géographique) mais par leur « proximité », notion déjà évoquée.

Dès lors, tandis que les scientomètres mesurent la similarité à partir du nombre normalisé de co-occurrences, les géographes ont tendance à intégrer des variables explicatives supplémentaires à leur modèle comme la distance physique (kilométrique) entre entités spatiales ou encore des variables qualitatives comme l'« effet pays » (l'appartenance nationale). La démarche est alors un peu différente de celle des scientomètres évoqués précédemment pour lesquels il s'agit de normaliser l'information à disposition avant de l'analyser. En analyse spatiale, l'approche consiste à estimer l'influence respective des facteurs explicatifs déterminants la force des interactions, et ensuite à se pencher sur la part de l'information n'ayant pu être expliquée par la taille et les déterminants spatiaux (les résidus).

À notre connaissance, les premiers qui ont essayé d'adapter le modèle gravitaire aux données de co-signatures pour estimer l'importance des contraintes spatiales sont des membres de la « *Swedish Regional Science Mafia* » (Beckmann, 1993). Au début des années 1990, la collaboration pionnière déjà évoquée entre le géographe Åke Andersson et le scientomètre Olle Persson est l'occasion d'expliquer la fréquence des co-publications internationales en tenant compte de l'effet conjugué de plusieurs facteurs (Å. E. Andersson & Persson, 1993). Les variables explicatives retenues par ordre d'importance sont : la taille scientifique (le nombre de publications par pays), le temps de voyage entre les pays, la proximité de langage, et la proximité politique (le régime). Un an après, et indépendamment, le scientomètre américain Sylvan J. Katz est le premier à estimer l'effet de la distance géographique sur les co-signatures entre universités à l'intérieur de plusieurs pays (Katz, 1994). Il montre que « la fréquence des collaborations scientifiques entre universités à l'intérieur du Royaume-Uni, du Canada et de l'Australie décroît exponentiellement avec la distance kilométrique entre partenaires. »⁸⁷ (*ibid.*). À la fin de son étude, il discute

⁸⁷ L'original : « the frequency of research collaboration between domestic universities in the United-Kingdom, Canada and Australia decreases exponentially with the distance separating the research partners. » (Katz, 1994).

de l'opportunité d'avoir recours à d'autres mesures de distance, considérant que la distance kilométrique correspond mal à la distance « vécue » entre universités ou villes appartenant à différents pays (en raison des discontinuités et de l'isotropie de l'espace géographique).

Au cours des années 2000, des efforts sont faits pour enrichir l'analyse quantitative des déterminants spatiaux des collaborations scientifiques. En particulier, les effets de « proximité institutionnelle » sont estimés en même temps que ceux de « distance géographique ». Ainsi, en utilisant un modèle de pseudo-régression, Nagpaul considère séparément l'effet des proximités géographiques, thématiques, et socio-économiques entre pays sur le taux de collaborations internationales par pays (Nagpaul, 2003). En 2009, Frenken *et al.* montrent à plusieurs échelles géographiques que la théorie de la « mort de la distance » (Morgan, 2001) ne se vérifie pas dans les pratiques de collaboration scientifique. En appliquant des équations gravitationnelles à trois jeux de données, ils montrent qu'en plus d'un effet de poids scientifique (effet de taille), la distance géographique (en kilomètres et temps de transport) et la distance « institutionnelle » (l'effet frontière) sont significatives pour expliquer l'intensité des collaborations scientifiques au cours des années 2000 premièrement entre 36 pays du monde, deuxièmement entre 1316 régions européennes (NUTS2), et troisièmement entre 40 régions néerlandaises (Frenken, Hoekman, *et al.*, 2009). Au-delà, ils montrent que l'effet de la distance a augmenté en même temps que diminuait l'effet frontière au sein de l'Union Européenne (Hoekman *et al.*, 2009). Compte tenu des politiques incitant au développement des collaborations scientifiques à l'échelle européenne, ils en déduisent qu'il s'agit d'une conséquence de l'élargissement de l'Union Européenne et d'une meilleure intégration des nations à cet espace. Enfin, depuis 2010, profitant des nouvelles méthodes de géolocalisation automatique des publications scientifiques, Tijssen *et al.* ont montré que la distance kilométrique de collaboration moyenne avait progressé globalement au cours des années 2000 tandis que la part des collaborations internationales s'était stabilisée (Tijssen, Waltman, & van Eck, 2012).

Cet apparent paradoxe sera discuté en détail dans les chapitres suivants. S'ils n'utilisent pas tous les mêmes variables et ne s'appuient pas sur les mêmes données où les mêmes périmètres d'étude pour le montrer, tous ces travaux reconnaissent le caractère déterminant de l'appartenance ou position géographique sur la participation à l'entreprise scientifique et l'armature géographique des collaborations. Cependant, on y reviendra, si certains sont convaincus que l'effet de la distance, qu'ils perçoivent comme un obstacle, est appelé à diminuer grâce à l'amélioration des moyens techniques de communication et de collaboration ; d'autres considèrent que la propension à collaborer plus ou moins souvent à longue distance dépend davantage de déterminants sociohistoriques que d'obstacles techniques. Comme l'indique le sociologue Göran Melin, il faut de toute façon avoir à

l'esprit que les facteurs structurels considérés au niveau macro (qu'ils soient techniques ou socio-historiques comme, par exemple, l'alliance ancestrale entre telle et telle université) ont peu de sens du point de vue des individus : « Lorsqu'on demande à un chercheur quelles sont les raisons pour lesquelles il collabore avec untel ou untel, il y a peu de chance qu'il mentionne l'amélioration des moyens de télécommunications, mais plutôt qu'un certain chercheur dispose de telles ou telles données ou qu'ils viennent juste de se rencontrer quelque part, se sont trouvés quelques intérêts communs et ont décidé de collaborer. »⁸⁸ (Melin, 2000).

Dans la littérature sur les sciences, le consensus n'existe pas non plus concernant l'effet de taille. Ni les causes géographiques, ni les conséquences géographiques de l'inégale répartition spatiale des activités scientifiques font l'unanimité. On sait que des pays très peuplés comme le Nigeria sont faiblement engagés dans l'entreprise scientifique quand d'autres, très petits comme la Suisse, Israël ou plus récemment Singapour et les Émirats-Arabis-Unis, misent beaucoup sur les activités de recherche. Comme l'ont très tôt constaté Robert K. Merton et Joseph Ben-David (voir Partie 1. Chapitre 4.1.1), la population d'un pays ne détermine pas mécaniquement sa participation à la production scientifique mondiale. D'autres facteurs importent au premier rang desquels l'investissement financier de l'État dans la R&D (Recherche et Développement) et la part du PIB consacré à la R&D, ce qui explique que ces indicateurs soient regardés de près avec d'autres facteurs socio-économiques par les institutions internationales comme l'UNESCO⁸⁹ (UNESCO, 2010). De même, lorsqu'on s'intéresse à la répartition des activités scientifiques à un niveau plus fin, on constate que ce ne sont pas les villes les plus peuplées qui accueillent nécessairement le plus de scientifiques. Il existe, en particulier dans les pays anglo-américains, des villes « universitaires » comme Saint-Andrews en Écosse, Ames (Iowa) aux États-Unis, Guelph au Canada qui sont des villes relativement peu peuplées, mais dont la dynamique principale relève de l'université. Il existe enfin d'importants centres de recherche localisés dans des lieux reculés comme le centre national d'Oak Ridge (Tennessee) aux États-Unis, qui doit son isolement à sa participation d'antan au projet Manhattan, on y reviendra.

Mais comme les grandes villes possèdent souvent « leur » université, on remarque que, sur l'ensemble des aires urbaines qui contribuent à l'activité scientifique, la population urbaine n'est pas complètement déconnectée de l'intensité de la production scienti-

⁸⁸ L'original : « If we were to ask a researcher about the reason for collaborating, it is unlikely that she/he would mention improved telecommunications, but rather that the collaborator had certain data or that they just met somewhere, realized that they shared an interest and decided that they could collaborate. » (Melin, 2000).

⁸⁹ Source: URL: <http://www.unesco.org/new/en/natural-sciences/science-technology/sti-policy/global-observatory-of-policy-instruments/>

fique. En effet, d'après le test que nous avons fait sur les données de production scientifique contemporaines, la population (nombre d'habitants) explique un peu plus de 35 % de la répartition de la production scientifique (articles, recensions et lettres du SCI Expanded en 2007, en moyenne mobile sur 3 ans) des 376 agglomérations scientifiques définies par notre équipe ($p\text{-value} = 2,2 \text{ e-}16$). Pour ce test, nous avons estimé la population des 376 aires urbaines en croisant leur surface avec les données de densité de population utilisées pour construire ces agglomérations (Center for International Earth Science Information Network - CIESIN - Columbia University *et al.*, 2005). La relation trouvée entre les deux variables est significative⁹⁰, mais certaines villes comme le Caire s'écartent notablement de la droite de régression linéaire, ce qui veut dire que cette capitale participe moins à la production scientifique mondiale que sa taille (en nombre d'habitants) pourrait le laisser imaginer. En fait, comme Andrea Bonaccorsi et Cinzia Daraio l'ont mis en évidence pour l'Italie et la France en 2005, la variable qui explique le mieux le niveau de production scientifique d'un lieu n'est pas sa population urbaine ou sa population étudiante, mais l'effectif de chercheurs qui y exercent (Bonaccorsi & Daraio, 2005). Plus récemment, un travail a trouvé des conclusions similaires pour le cas français (les 20 agglomérations françaises qui publient le plus) (Grossetti, Eckert, Maisonobe, & Tallec, à paraître). La prédiction établie est très bonne puisque le coefficient de corrélation multiple (R^2) donnant une idée de la part de variation expliquée par le modèle est cette fois-ci de 95 %. Une étude plus systématique de ce rapport entre effectif de chercheurs et volume de production mériterait d'être conduite à l'échelle du monde mais un lourd travail de recueil des données et de mise en adéquation des effectifs de chercheurs avec les périmètres d'étude (les agglomérations) devrait alors être réalisé — ce qui explique que notre test soit limité au caractère explicatif du nombre d'habitants des agglomérations sur leur production scientifique.

Ces résultats, même partiels, suggèrent que l'effet de taille, ou le fait que les entités spatiales participent inégalement à la production scientifique mondiale, est la conséquence du fait que les chercheurs sont inégalement répartis dans l'espace géographique. Or, à l'encontre de cette conception, les tenants de la théorie de la « masse critique » considèrent que l'effet de taille n'est pas « purement structurel ». Selon eux, l'inégale répartition géographique de la production scientifique serait la conséquence de l'inégale capacité des lieux à contribuer à l'entreprise scientifique : certains lieux seraient plus propices que d'autres à la pratique scientifique. Cette considération, déjà évoquée (voir Partie

⁹⁰ Comme Grossetti et Tallec, j'ai considéré le logarithme des variables avant de leur appliquer une régression linéaire. En effet, ces variables ayant une distribution de type log-normal, l'application à toutes les variables de la fonction logarithme permet de se ramener à des lois normales et d'améliorer ainsi la qualité du modèle.

1. Chapitre 3.2), a été beaucoup mobilisée comme argument pour s'opposer aux politiques de décentralisation de la recherche. D'après la théorie de la « masse critique », les chercheurs produiraient d'autant plus qu'ils sont regroupés aux mêmes endroits. Ce surplus de performance associé à la concentration des chercheurs en un même lieu serait de nature à expliquer l'inégale répartition des chercheurs dans l'espace. En effet, pour être plus performants, les chercheurs chercheraient spontanément à se regrouper. Dans cette perspective, l'effet de taille serait auto-entretenu. Or, contrairement à ce que suppose cette théorie, la dispersion spatiale des chercheurs et, par ricochet, la dispersion de la production scientifique dans l'espace géographique a été continue depuis les années 1960-1970. Ainsi, en 1977, Inhaber a observé les premières traces de la déconcentration de la production scientifique entre villes à l'échelle mondiale (Inhaber, 1977). Puis, en 1991, Brocard a montré qu'à l'échelle de la France cette déconcentration n'était pas seulement le résultat des politiques de décentralisation de l'enseignement supérieur, dans la mesure où les effectifs de chercheurs CNRS au niveau régional s'étaient progressivement alignés sur les effectifs des enseignants-chercheurs des universités : entre 1964 et 1984 le coefficient de corrélation linéaire entre les deux populations serait passé de 0,67 à 0,89 (Ile-de-France exclue) (Brocard, 1991, p. 74).

Dans l'ensemble, tout ceci suggère que l'inégale répartition de la production scientifique n'est le reflet ni d'une inégale performance scientifique des chercheurs imputable à leur lieu d'exercice, ni de l'attraction de certains lieux au détriment d'autres en raison de la taille ; mais, plutôt, de l'inégale répartition des forces de travail scientifique entre les lieux qui, aidée par la dynamique de décentralisation de l'enseignement supérieur et des organismes de recherche nationaux, n'aurait cessé de s'atténuer depuis les années 1960-1970. En fait, comme l'a expliqué Pumain en discutant de la théorie des « avantages d'agglomération », il n'est pas nécessaire de faire l'hypothèse que la concentration des activités améliore les performances et la productivité pour expliquer la concentration des activités en quelques lieux, une théorie évolutive suffit (Pumain, 1997). Les déterminants de la localisation des activités sont forcément multiples et, dans cette quête, les études de cas présentent, selon nous, un grand intérêt parce qu'elles permettent la prise en considération des facteurs historiques et donc des effets d'inertie.

L'idée d'une meilleure performance des lieux où l'activité scientifique est la plus concentrée n'est pas la seule idée circulant à propos des conséquences de l'effet de taille. Une autre conviction circule, qui ne concerne pas la productivité scientifique, mais l'organisation des échanges scientifiques. Dès 1979, Frame et Carpenter montrent que le fait d'exercer dans un grand ou un petit pays (en termes de production scientifique) modifie la propension des scientifiques à collaborer plus ou moins avec l'étranger (Frame & Carpenter, 1979). En effet, ils établissent une relation entre la taille des pays (en nombre

de publications) et la part des relations internationales. Cette dernière apparaît d'autant plus élevée que le pays participe peu à la production scientifique mondiale. L'interprétation est que les chercheurs appartenant à des petits pays sont davantage poussés à collaborer avec l'étranger dans la mesure où ils ont moins de partenaires avec qui travailler à l'intérieur des frontières de leur pays. Cela suppose une fois de plus que l'effet de taille n'est pas neutre et purement structurel, mais en mesure d'influencer les pratiques des chercheurs au niveau individuel. À l'encontre de cette théorie, Göran Melin s'est associé à Olle Persson, à la fin des années 1990, pour montrer que les pratiques individuelles de collaboration n'étaient pas différentes selon la taille des lieux, au moins dans le cas des universités. Une première étude limitée aux universités scandinaves a montré que la taille des universités n'influait pas le volume et la composition des collaborations des chercheurs (Melin & Persson, 1998). Puis, une seconde étude a été conduite par Melin, seul, qui a trouvé des résultats comparables pour un groupe d'universités américaines, ce qui l'a conduit à remettre en cause la relation établie par Frame et Carpenter entre la taille scientifique des pays et l'intensité de leurs collaborations internationales (Melin, 1999). Selon Melin, la propension individuelle à collaborer ne varie pas en fonction de la taille des pays mais plutôt en fonction de l'ouverture culturelle et de la politique des pays en matière scientifique. En effet, les résultats de Melin indiquent que la propension des chercheurs à collaborer à l'international dépend davantage de facteurs institutionnels comme la culture, la langue, et donc le pays d'appartenance, que d'invariants comme la taille scientifique de leur université ou de leur pays.

Ces résultats suggèrent que la répartition de la production scientifique ne peut suffire à elle seule à interpréter ou prévoir les évolutions et la structuration de la géographie de la production scientifique. Il faut tenir compte de variables extérieures comme la diffusion et la multiplication des établissements d'enseignement supérieur pour comprendre la diffusion géographique et l'augmentation des chercheurs dans les villes universitaires et, au-delà, la déconcentration géographique de la production scientifique. De même, il faut tenir compte de la nationalité des chercheurs et des spécificités de leur pays d'appartenance pour comprendre leur plus ou moins grande propension à collaborer avec des chercheurs étrangers.

Enfin, en plus de la taille et de la similarité, une dernière variable interne à l'activité scientifique est fondamentale pour les scientomètres : la discipline scientifique. Un certain nombre de travaux discutent des moyens de normaliser les indicateurs scientométriques, tel que le facteur d'impact d'une revue, afin de les rendre comparables d'une discipline à l'autre, d'une institution à l'autre, d'une revue à l'autre etc. (Glänzel, Thijs, Schubert, & Debackere, 2009 ; Opthof & Leydesdorff, 2010 ; van Raan, van Leeuwen, Visser, van Eck, & Waltman, 2010 ; Gingras & Larivière, 2011 ; Waltman, van Eck, van

Leeuwen, Visser, & Van Raan, 2011 ; Opthof & Leydesdorff, 2011). En effet, selon les disciplines, les pratiques de production, de collaboration et de citation sont différentes. Au sein même de la recherche biomédicale, il existe d'importantes différences en matière de citation. Un article récent montre en particulier que les recherches les plus fondamentales ont un meilleur impact que les recherches cliniques (van Eck, Waltman, Van Raan, Klautz, & Peul, 2013). Aussi, les différences en termes de collaboration sont reflétées par l'amplitude variable du nombre de co-auteurs par article selon les domaines, le degré de collectivisation tendant à augmenter avec la « lourdeur des équipements requis (physique des particules), la complexité et la diversité des tâches à accomplir (médecine clinique), et l'échelle de la recherche (océanographie, espace) » (Gingras, 2002). Ainsi, les productions en sciences de l'univers peuvent être co-signées par un millier de partenaires à la fois tandis que les productions en mathématiques ou en sciences humaines sont rarement co-signées par plus de 2 ou 3 partenaires. Dès lors, pour saisir l'évolution de la géographie des activités de production scientifique toutes disciplines confondues, comme c'est le cas dans le chapitre suivant (Chapitre 9), la méthode de comptage bibliométrique que nous avons retenue est intéressante car elle ne traite pas toutes les co-signatures sur un pied d'égalité. Elle normalise la valeur des collaborations en fonction du nombre de partenaires par co-publication.

Dans les chapitres suivants, nous évoquerons de nombreuses variables explicatives sans pour autant les intégrer systématiquement à des modèles mathématiques permettant d'établir des causalités. En ce sens, l'approche que nous allons privilégier sera plus souvent « narrative » qu'explicative. Pour Andrew Abbott, qui défend le recours aux méthodes narratives : « il n'y a aucune raison de penser qu'un indicateur donné ait un quelconque statut particulier, expliqué ou explicatif » puisque « la vie sociale est une immense toile dans laquelle tout se tient » (Abbott, 2011). Les méthodes standards cherchant à établir des causalités parviennent difficilement à considérer simultanément toutes les dépendances et sont bornées par l'hypothèse de l'individualisme méthodologique. Dans ces modèles, « la trajectoire causale serait toujours la même » or « la signification des causalités dépend du contexte » et « c'est la succession des points, envisagés dans toutes les dimensions à la fois, qui importe » (*ibid.*). La méthode narrative proposée par Andrew Abbott, est une approche descriptive et non causale : « Dans ce cadre, étudier une succession d'étapes n'implique pas de poser des questions de causalité, mais plutôt de descriptibilité » (*ibid.*). Étant donnée l'échelle à laquelle nous travaillons et le nombre d'observations que nous avons à considérer, cette approche est adaptée car elle est en mesure de nous aider à saisir l'ensemble des observations en un seul regard ainsi qu'à mettre en évidence la dépendance des observations à leur position spatio-temporelle.

CHAPITRE 9. LES ÉVOLUTIONS DE LA GÉOGRAPHIE DES ACTIVITÉS SCIENTIFIQUES CONTEMPORAINES

Le WoS n'est rien d'autre qu'un ensemble de notices bibliographiques mais c'est avec *Scopus* l'un de ceux qui en compte le plus, soit entre 1 et 1,5 millions de publications scientifiques par an au cours des années 2000 (pour 500 000 en 1987). Les données *Scopus* sont assez comparables, et des estimations réalisées à partir des chiffres de 2013 suggèrent qu'il y aurait désormais dans le monde pas moins de 4 millions d'auteurs produisant plus de 2 millions d'articles visibles par an (Plume & van Weijen, 2014). Comme nous avons établi précédemment que l'ensemble du WoS était révélateur d'un certain « monde scientifique », celui de la « science visible sous influence américaine », il est pertinent d'en explorer le contenu à plusieurs dates. Cependant, avant 2008, les enregistrements du WoS ne contiennent pas d'indications permettant de connecter les auteurs à leurs adresses. Autrement dit, pour chaque publication, on dispose de la liste des auteurs et de la liste des adresses, mais on ne sait pas à quels auteurs correspondent quelles adresses. Cette caractéristique est importante car elle élimine la possibilité de localiser les auteurs des publications écrites en collaboration entre plusieurs lieux. Le lien entre auteurs et adresses est d'autant plus difficile à établir que les auteurs peuvent avoir plusieurs affiliations, c'est-à-dire qu'ils peuvent appartenir simultanément à plusieurs organismes de recherche. De ce fait, même une publication signée par un seul auteur est susceptible d'être multi localisée. Ces subtilités ont une incidence sur les traitements géographiques qu'il est possible d'envisager à partir d'un *corpus* de données bibliographiques⁹¹.

Pour exploiter l'information géographique du WoS pour plusieurs séries d'années antérieures à 2008 et ainsi obtenir des « photographies » de l'état de la production mondiale contemporaine à différents moments, la méthode mise au point collectivement avec les membres de la « tâche 2 » du programme de recherche « Géoscience » se concentre

⁹¹ Évidemment, rien n'empêche, pour mieux localiser les auteurs et si la recherche le justifie, de coupler l'information bibliographique avec une information biographique. Il faut aussi savoir que la situation a changé pour les publications indexées depuis 2008 puisque des connecteurs permettent désormais d'associer chaque auteur à son ou ses adresse(s) (voir l'exemple de la Figure p. 192). À l'opposé, pour les publications parues avant les années 1990, il y a rarement plus d'une adresse renseignée par publication, y compris quand il s'agit d'articles écrits en collaboration entre plusieurs lieux (parce qu'il était alors fréquent de ne retenir que la première adresse pour l'indexation), mais nous y reviendrons encore plus tard.

sur la localisation des publications à travers le traitement des adresses et ignore les auteurs. S'il se justifie par la limite que nous venons d'évoquer, ce parti pris possède un fondement théorique : bien que largement corrélées, la géographie de la production et la géographie des producteurs sont deux choses différentes et c'est la première que nous étudions. Le WoS est une base de publications et la philosophie avec laquelle la base a été conçue consiste à envisager « la science » comme « l'ensemble des publications scientifiques ». Dès lors, la géographie qu'elle permet avant tout de considérer est celle de la production. En suivant cette voie, nous acceptons implicitement de considérer à la suite de Derek de Solla Price, le père de la scientométrie, que les productions scientifiques sont un bon « indicateur social » et qu'elles peuvent nous renseigner sur l'organisation de « la science » (Solla Price, 1962 cité par Polanco, 1995).

1. Géographie de la production scientifique contemporaine

*« Le Monde est système, et le système des acteurs de son espace est mondial ; mais il fonctionne à plusieurs niveaux géographiques, entre lesquels abondent les contradictions. »
(R. Brunet, 2001, p. 59)*

1.1. Une approche multi-niveaux pour saisir la croissance de la science

Prises dans leur intégralité, les données du WoS permettent de mettre en évidence le phénomène de « croissance de la science », c'est-à-dire le fait qu'un nombre croissant de chercheurs produisent un nombre croissant de publications depuis un nombre croissant de lieux. Avant la mise au point par notre équipe de la méthode de « scientométrie spatiale » précédemment présentée, il n'existait pas de travaux qui soient arrivés à spatialiser ce phénomène à un niveau aussi fin que le niveau urbain et à une échelle aussi petite que l'échelle mondiale.

Cependant, une bonne connaissance des événements intervenus à partir des années 1940 dans plusieurs pays pouvait laisser deviner comment les choses allaient tourner. En effet, après la Seconde Guerre mondiale, le développement des activités scientifiques a été encouragé dans de nombreux pays à la fois pour des raisons géopolitiques et socio-économiques. Le coût matériel et humain associé à certaines recherches de pointe a considérablement augmenté, justifiant des investissements accrus de la part des États mais aussi

l'embauche d'un nombre croissant de professionnels. L'élévation du niveau général d'éducation de la population a suscité une demande en enseignants-chercheurs, lesquels ont été amenés à faire fonctionner des universités de plus en plus peuplées et de mieux en mieux réparties sur les territoires. Ainsi serions-nous passés, d'après John Meyer, de 500 000 étudiants dans le cycle tertiaire dans le monde en 1900 à plus de cent millions en 2000. Cette expansion ne s'est pas limitée aux pays « riches et industrialisés », mais aurait aussi participé à la « globalisation » en facilitant « le mouvement de professionnels et le développement d'une culture mondiale commune » (Schofer & Meyer, 2005). Ce phénomène de croissance a naturellement suscité des interrogations en matière d'organisation des scientifiques. En dehors des arguments humanistes et pacifistes pour coordonner les activités scientifiques au niveau international⁹², des arguments pragmatiques liés à la « croissance de la science » ont volontiers été mobilisés. En particulier, le géographe Jean Gottmann en a usé en 1973, lorsqu'il a été invité à proposer un papier au *Nobel Symposium* d'Oslo sur le besoin d'une politique internationale de la science (Gottmann, 1973). Alors qu'une bonne partie des projets de laboratoires internationaux imaginés dans l'immédiat après-guerre au sein de l'ONU, notamment par Gottmann, avait échoué (Petitjean, 2001), ce dernier considérait que le besoin d'une coordination des chercheurs n'avait pas disparu pour autant : « La dépendance grandissante d'une proportion croissante de l'humanité envers les sciences et technologies pour des besoins pratiques quotidiens est souvent associée avec la métaphore d'un "monde rétrécissant". Il semble, en fait, que le progrès scientifique et technologique participe d'un élargissement du monde ; chacun de nous, comme individus, ayant à composer avec davantage de gens, dans davantage d'endroits, pour récolter les données, les informations et les conseils nécessaires ; chaque pays doit aussi traiter avec un nombre croissant de lieux, d'individus, d'institutions, et de concepts. Cette expansion de notre univers privé et public suppose davantage d'ordre, de rangements, et de lignes directrices. C'est pourquoi, bien que les scientifiques s'opposent à toutes tentatives de gouverner la science mondiale, ils ne peuvent qu'accueillir la mise en place de dispositifs qui puissent faciliter l'accès à leur besoins et ressources. Le temps est d'autant plus propice au développement d'une politique internationale que l'ordre apparent qui existait après 1945, dû à la forte concentration des activités scientifiques dans un petit nombre de pays et de centres, s'est érodé. Si une politique internationale des activités scientifiques a été opérante au cours de ces 20 dernières années, c'est certainement celle de la décentralisation géographique. Elle a connu des résultats ; l'un d'entre eux est précisément le besoin accentué de normes communes. »⁹³ (Gottmann, *op. cit.*).

⁹² Voir la rhétorique de l'« internationalisme scientifique », évoquée dans la Partie 1, Chapitre 2.

⁹³ L'original : « The growing dependence of a higher proportion of mankind on science and technology for daily critical requirements is sometimes associated with the metaphor of a 'shrinking world'. It seems,

Ainsi, tant que le nombre de chercheurs susceptibles de communiquer et travailler ensemble et que le nombre de lieux où ils étaient accueillis restaient limités, le paysage de l'activité scientifique demeurait lisible. À partir du moment où les activités scientifiques ont été décentralisées au sein d'un nombre croissant de pays et qu'il a fallu simultanément veiller à ce que les chercheurs s'intègrent à de grands ensembles mondiaux, les défis à relever ont été de plusieurs ordres :

- d'ordre politique : comment administrer et répartir les moyens entre davantage de sites et veiller à maintenir une certaine cohérence entre les différentes strates administratives ?

- d'ordre scientifique : comment s'informer de ce qu'il se passe dans un nombre croissant de sites, communiquer et travailler avec la part de plus en plus grande de chercheurs investissant de nouveaux espaces d'activité ?

- et d'ordre analytique : comment saisir simultanément les logiques infra-, inter- et supra- nationales de croissance et de structuration des activités scientifiques ?

Durant la guerre froide, la lecture du monde est axée sur l'opposition entre un bloc de pays associés aux États-Unis, un autre bloc associé à l'URSS et un dernier bloc de pays réclamant une autonomie à l'égard des grandes puissances : le « Tiers-Monde ». À cette période, le niveau national apparaît comme le niveau « naturel » d'analyse pour étudier l'organisation du système scientifique mondial (Beaver & Rosen, 1978 ; Frame & Carpenter, 1979 ; Schubert & Braun, 1990 ; Okubo, Miquel, Frigoletto, & Doré, 1992 ; Luukkonen, Tijssen, Persson, & Siverstsen, 1993). Mais, depuis le début des années 1990, sans toutefois que cette approche disparaisse (Schubert & Glänzel, 2006), il est devenu fréquent de substituer le terme « globalisation » à celui d'« internationalisation » des activités scientifiques tandis que les niveaux d'organisation supra- et infra- nationaux ont commencé à revendiquer une place dans la structuration des activités scientifiques. À partir de là, l'hypothèse d'une tension entre la « régionalisation » de la recherche (Brocard, 1991) et sa « globalisation » (Schott, 1993) a été envisagée, à laquelle on peut ajouter le thème de la « continentalisation » ou de l'organisation des activités scientifiques à l'échelle

in fact, that scientific and technological work proceeds in an expanding world; each of us, as an individual, has to deal with a constantly growing number of people, in more places, to gather the data, information, and advice we need; every country deals with a growing number of places, people, institutions, and concepts. This expansion of our private and public universe requires more ordering, classification, guidelines. Scientists would certainly resist any attempt to marshal world science; they might welcome the establishment of a recognized marshalling yard for needs and resources. Last but not least, the time seems ripe for an international policy because the apparent order which existed after 1945, owing to the high concentration of science and technology in a very small number of countries and centre, has now been eroded. If any international policy for science has been operating in the last 20 years, it has certainly been one of geographic decentralization. It has borne results; one consequence of it is precisely the need for more common moorings. » (Gottmann, 1973).

de grandes aires géographiques ou culturelles (M. Leclerc & Gagné, 1994 ; Melin & Persson, 1998).

Pour saisir et interroger ces phénomènes, la distinction entre intra- et international est devenue insuffisante, ce qui justifie parfois l'emploi du terme « transnational » puisque des échanges peuvent se mettre en place entre scientifiques, qui ne sont pas médiés par l'instance nationale et/ou dont la logique ne cadre pas avec la tendance nationale (Crawford, Shinn, & Sörlin, 1993). Notamment, l'idée de « globalisation » suppose que des logiques « globales » puissent influencer l'évolution du système indépendamment des patrons nationaux. Dans ce registre, plusieurs théories co-existent. Certains chercheurs présument que la globalisation se fait au prix d'une concentration des activités au sein d'un nombre limité de pôles à vocation mondiale, tandis que d'autres considèrent que la globalisation est un phénomène capable d'effacer la hiérarchie entre les lieux, voire l'importance de la géographie — c'est la « mort de la distance » (Morgan, 2001), expression attribuée à Frances Cairncross. Tenante de la première approche, la spécialiste en politiques de la science Caroline Wagner considère que : « En dépit de l'accélération de la diffusion des données scientifiques, de l'information et de la connaissance, le monde de la science est toujours loin d'être plat. Mais la focale est passée du niveau national au niveau global. Les réseaux auto-organisés qui tapissent le globe aujourd'hui en sont la caractéristique actuelle la plus notable. Ces réseaux forment un collège invisible de chercheurs qui collaborent non pas parce qu'on leur a dit de le faire mais parce qu'ils ont envie de le faire, qui travaillent ensemble non pas parce qu'ils partagent un laboratoire ou une discipline mais parce qu'ils peuvent s'apporter des idées, des connaissances et des compétences complémentaires (...). Au 21^e siècle, le *melting pot* des scientifiques, l'appartenance nationale et les allégeances jouent un rôle mineur. La curiosité scientifique et l'ambition sont les forces majeures en place dans le nouveau collège invisible. »⁹⁴ (Wagner, 2008, p. 2). Aussi, dans sa thèse sur la globalisation de la science, qui traite entre autres de la standardisation des essais cliniques, Jarno Hoekman évoque cette approche individualiste. En se fondant sur le travail d'Anna Lee Saxenian, il explique : « Comme les barrières spatiales à la collaborations ont disparu, les scientifiques sont de plus en plus capables d'auto-organiser leurs collaborations gardant en tête leurs intérêts personnels. La mobilité crois-

⁹⁴ L'original : « Despite the accelerating diffusion of scientific data, information, and knowledge, the world of science remains far from flat. But its focus has changed from the national to the global level. Self-organizing networks that span the globe are the most notable feature of science today. These networks constitute an invisible college of researchers who collaborate not because they are told to but because they want to, who work together not because they share a laboratory or even a discipline but because they can offer each other complementary insight, knowledge, or skills (...) In the twenty-first century melting pot of science, national citizenship or allegiance plays a minor role. Scientific curiosity and ambition are the principal forces at work in the new invisible college. » (Wagner, 2008, p. 2).

sante peut dans ce cas entraîner une concentration accrue des chercheurs réputés dans quelques grands centres au lieu d'une répartition plus équitable de l'activité scientifique à travers le globe. »⁹⁵ (Saxenian, 2006 citée par Hoekman, 2012, p. 22).

Tandis que la plupart des « globalistes » se focalisent sur l'accentuation des concentrations et des hiérarchies spatiales en matière d'activités scientifiques pour contrer l'idée d'une l'homogénéisation ou d'un « aplatissement » du monde (le débat entre Richard Florida et Thomas L. Friedman, déjà évoqué, est un exemple populaire de cette opposition appliquée aux activités économiques, voir Partie 1.Chapitre 3.2.2), une troisième voie se dessine. Cette troisième voie permet d'insister sur le rôle structurant des organisations à au moins trois niveaux géographiques : local, national, mondial. Selon cette approche, l'individu n'est pas tout seul face au monde mais il est la conjugaison de ses appartenances institutionnelles en un temps et un lieu donné. Parmi les appartenances, nombreux sont les spécialistes des activités scientifiques qui insistent sur la prégnance de l'appartenance nationale. En effet, comme le suggère Luke Georghiou à l'issue d'une réflexion sur l'évolution des collaborations scientifiques au cours des années 1980-1990 : « aussi longtemps que l'État-Nation reste une unité compétitive, il y a des chances que la coopération internationale continue d'être perçue comme un moyen d'élever la position de la science nationale plutôt que comme un moyen de la remplacer. »⁹⁶ (Georghiou, 1998). Dans le même esprit, Melin va jusqu'à remettre en cause l'idée de « globalisation » car il remarque que les institutions scientifiques ont toujours tendance à adopter un comportement collaboratif caractéristique de leur pays d'appartenance (il étudie le comportement d'une population d'universités) (Melin, 1999).

À notre sens, même si les États restent structurants, il est intéressant de s'autoriser une réflexion à plusieurs niveaux géographiques, et c'est d'ailleurs la meilleure façon d'estimer la progression du poids du niveau national par rapport aux autres niveaux d'organisation. De plus, en raisonnant à plusieurs niveaux, il devient possible d'envisager de concert des phénomènes qui seraient contradictoires autrement. Ainsi, pour Terry Shinn, il n'est pas question de nier la globalisation mais de montrer qu'elle n'annule pas le rôle structurant des États en matière de sciences et techniques : « Il est indispensable de s'éloigner du scénario "soit, soit" qui met l'accent sur les pratiques et les configurations soit principalement globales, soit principalement nationales, les unes excluant les autres. »

⁹⁵ L'original : « As spatial barriers to collaborate vanish, scientists are increasingly able to self-organize their collaborations keeping in mind their personal interests. Increased mobility can in this case for instance result in a concentration of reputable researchers in a few scientific centers rather than in a more equal distribution of scientific activity across the globe. » (Saxenian citée par Hoekman, 2012, p. 22).

⁹⁶ L'original : « so long as the nation-state is a competitive unit, it is likely that international cooperation will be seen as a means of enhancing the position of national science base rather than replacing it. » (Georghiou 1998).

(Shinn, 2002). Au cours des années 2000, Rigas Arvanitis *et al.* prolongent et intègrent la question régionale à cette réflexion multi-niveaux : « Dans de nombreux pays, les politiques de recherche et d'innovation tentent de concilier une intégration de la science nationale dans de grands ensembles mondiaux et une décentralisation régionale des activités de recherche. Pour examiner ces dynamiques, il faut réfléchir simultanément à l'échelle mondiale et régionale. » (Arvanitis *et al.*, 2008). C'est ce que nous permet de faire la méthode présentée dans les chapitres précédents puisqu'en repérant l'évolution du volume d'activité scientifique au niveau des villes, elle permet de mettre en relation la progression locale de cette activité avec la progression de l'activité mesurable au niveau national et au niveau mondial. De cette façon, il est possible de détecter les différentes manifestations d'une même observation selon qu'elle est regardée à une échelle ou une autre. Au-delà, il est possible de voir comment le réseau des collaborations scientifiques est impacté par l'évolution de la géographie de la production scientifique. Cette approche permet d'éprouver et de généraliser les résultats obtenus dès le début des années 2000 et encore aujourd'hui à travers des études de cas sur l'évolution des activités scientifiques au sein de plusieurs pays.

1.2. La déconcentration des activités scientifiques à plusieurs niveaux

Dès le début des années 2000, en spatialisant les données de production scientifique à l'échelle française, Michel Grossetti et Béatrice Milard notent, sans nier le processus de croissance des collaborations scientifiques internationales, un développement des collaborations locales (intra-départementales ou -régionales). Cette évolution semble aller de pair avec la déconcentration spatiale des activités d'enseignement supérieur et de recherche. Il advient rapidement que cette déconcentration n'est pas limitée à la France mais a également lieu dans d'autres pays tels que l'Espagne, le Portugal, l'Afrique du Sud, ou encore la Russie (Grossetti & Losego, 2003 ; Milard & Grossetti, 2006 ; Grossetti *et al.*, 2009). Ce phénomène résulterait de la construction et de la densification en plusieurs étapes des systèmes nationaux de recherche de ces différents pays. Une typologie s'adaptant à tous les cas de figure étudiés se révèle opérante qui distingue : les « sites-capitales », les « autres sites dont les établissements d'enseignement supérieur et de recherche ont été créés avant les années 60-70 et qui se présentent généralement comme des pôles d'envergure "nationale" » ; les « sites créés à la suite de la massification des années 60-70 qui sont aujourd'hui plutôt des pôles "régionaux" » ; et les « sites créés après la seconde massification, celle des années 80-90, et qui constituent souvent des sites d'aménagement des territoires intra-régionaux » (*ibid.*, 2009).

Entre ces différents sites, les données de production scientifique extraites du SCI Expanded – l'index bibliographique centré sur les sciences de la nature et de la technique jugé le plus représentatif de la science internationale – témoignent d'un rééquilibrage. La part qu'occupent les sites secondaires dans la production scientifique de ces différents pays au cours de la dernière décennie est en effet de plus en plus importante (Tableau 18).

	Nb. circ.	1 ^{ère} période	2 ^{ème} période	
FRANCE				
		1990-92	1998-00	
Centres scientifiques régionaux	23	85,7%	84,5%	↘
Autres départements des régions	65	14,3%	15,5%	↗
Total	88	100%	100%	
ESPAGNE				
		1990-92	1998-00	
Centres scientifiques régionaux	12	80,7%	75,1%	↘↘
Autres provinces des régions	31	19,3%	24,9%	↗↗
Total	43	100%	100%	
PORTUGAL				
		1990-92	1998-00	
Centres scientifiques régionaux	3	84,4%	69,8%	↘↘↘
Autre districts des régions	10	15,6%	30,2%	↗↗↗
Total	13	100%	100%	
AFRIQUE DU SUD				
		1989-93	2001-03	
Centres universitaires provinciaux		75.3%	65.2%	↘↘
Autres universités des provinces		24.7%	34.8%	↗↗
Total		100%	100%	
RUSSIE				
		1992-94	2001-03	
Centres scientifiques de la province ¹	6	67,6%	64,5%	↘
Autres sites	64	32,4%	35,5%	↗
Total				

Tableau 18 – La déconcentration intra-régionale des publications scientifiques (1990-2000). Tableau extrait du chapitre « La territorialisation comme contrepoint à l'internationalisation des activités scientifiques » (Grossetti, Losego & Milard, 2009)

En accord avec ce constat, d'autres monographies de pays montrent une tendance à la déconcentration intra-nationale de l'activité scientifique au détriment des régions capitales en Chine (D. E. Andersson, Gunessee, Matthiessen, & Find, 2014 ; Ma, Fang, Pang, & Li, 2014) et en Corée du Sud (Shapiro, So, & Park, 2009). Mais il faut considérer l'évolution de la répartition de l'ensemble des publications (articles, recensions et lettres) du SCI Expanded sur la période récente pour estimer à quel point les agglomérations les plus publiantes du monde ont perdu du poids dans la production scientifique à plusieurs échelles.

Pour commencer, le poids dans la production scientifique nationale respective des trente centres de production scientifique les plus importants du monde a diminué par rapport à leur poids dans les années 1980. D'après le Tableau 19, cette baisse a particulière-

ment touché les centres chinois (Beijing, Hong Kong et Shanghai) entre 2000 et 2007. Au cours de cette période, le poids de Séoul a légèrement augmenté, mais ce regain succède à une forte déconcentration intervenue entre 1987 et 2000 pendant laquelle le poids de Séoul est passé de 83 % à seulement 52 % de la production du pays.

Agglomérations les plus productives en 2007*	Poids dans la production nationale (%)			Evolution (%2000*-%2007*)	Nombre de publications en 2007*
	1987*	2000*	2007*		
Tokyo	34,1	31,9	33,3	1,4	24787,0
Beijing	34,4	25,3	21,1	-4,1	20394,2
Paris	45,2	36,9	34,7	-2,2	16195,3
New York	7,2	6,5	5,9	-0,6	16087,0
Seoul	82,6	52,5	55,4	2,9	15859,5
Boston	5,2	5,7	5,6	-0,1	15196,3
Londres	27,1	25,2	24,0	-1,2	14935,1
Kyoto-Osaka	21,7	19,7	19,2	-0,4	14307,8
Baie de San-Francisco	5,4	5,0	4,7	-0,2	12844,6
Washington-Bethesda	5,5	5,0	4,6	-0,5	12413,3
Los Angeles	4,7	4,2	4,5	0,3	12266,7
Shanghai	18,4	10,9	10,3	-0,7	9929,4
Moscou	57,6	44,4	42,4	-2,0	9262,3
Taipei	51,4	51,3	44,4	-6,9	8797,1
Chicago	3,0	3,0	2,9	-0,1	7854,6
Philadelphie	2,9	3,0	2,8	-0,3	7503,0
Toronto	19,6	20,5	20,0	-0,4	7502,5
Madrid	31,1	26,3	23,5	-2,8	6937,4
Durham Research Triangle	2,0	2,5	2,5	0,0	6757,7
Berlin	10,0	9,7	10,1	0,4	6403,5
Hong Kong	--	14,5	6,5	-7,9	6306,2
Sydney	24,1	26,1	26,3	0,3	6255,2
Singapour	100,0	100,0	100,0	0,0	6133,6
Milan-Pavie	--	15,9	14,7	-1,2	5819,9
Rome	15,0	14,9	14,6	-0,2	5787,3
Munich	9,7	9,3	9,0	-0,3	5731,1
Montreal	13,8	15,4	15,2	-0,2	5696,4
Melbourne	24,6	23,3	24,0	0,6	5695,8
Baltimore	1,7	2,1	2,0	0,0	5588,4
Barcelone	19,8	19,2	18,7	-0,5	5525,6

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: SCIEp (articles, recensions et lettres)

Tableau 19 – L'évolution du poids des grands centres de production scientifique au sein de leurs espaces nationaux de 1987 à 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Au regard des monographies existantes, cette régression se comprend comme un phénomène de rattrapage des sites secondaires – créés ou renforcés à l'occasion des politiques de décentralisation de la recherche et de démocratisation de l'enseignement supérieur – vis-à-vis des pôles principaux des pays les plus actifs du système scientifique mondial. Pour autant, il se pourrait que comme l'affirment Matthiessen *et al.*, ces pôles majeurs puissent, malgré la baisse de leur poids relatif à l'échelle de leur propre pays, si-

multanément accroître leur caractère nodal ou polarisant à l'échelle mondiale (Matthiesen *et al.*, 2010). Or, en rapportant la production de ces « villes mondiales » à la production mondiale en 2000 et en 2007 (Tableau 20), on constate qu'à l'exception des centres chinois, tous les centres de production les plus importants du monde pèsent de moins en moins lourd dans la production scientifique mondiale ainsi que dans le sous-ensemble des co-publications entre agglomérations — ce qui contredit les représentations de Matthiesen *et al.* (Grossetti *et al.*, 2014).

Agglomérations les plus publiantes en 2007*	Poids dans la production mondiale	Poids dans les co-publications mondiales	Nombre de publications en 2007*	Nombre de co-publications en 2007*
	Evolution relative (%2000*-%2007*)			
Tokyo	-0,6	-0,5	24787,0	6943,5
Beijing	0,9	0,8	20394,2	5899,5
Paris	-0,4	-0,4	16195,3	5705,5
New York	-0,4	-0,3	16087,0	5582,5
Seoul	0,6	0,3	15859,5	3269,4
Boston	-0,2	-0,1	15196,3	5542,9
Londres	-0,5	-0,2	14935,1	5115,5
Kyoto-Osaka	-0,4	-0,3	14307,8	4211,8
Baie de San-Francisco	-0,2	-0,3	12844,6	4536,0
Washington-Bethesda	-0,3	-0,3	12413,3	4948,7
Los Angeles	-0,1	-0,1	12266,7	4060,2
Shanghai	0,5	0,4	9929,4	2639,8
Moscou	-0,4	-0,3	9262,3	2077,5
Taipei	0,2	0,2	8797,1	2091,5
Chicago	-0,1	-0,1	7854,6	2670,8
Philadelphie	-0,2	-0,1	7503,0	2554,7
Toronto	0,0	0,0	7502,5	2358,6
Madrid	0,0	0,0	6937,4	2029,6
Durham Research Triangle	-0,1	0,0	6757,7	2515,2
Berlin	-0,1	-0,1	6403,5	2584,0
Hong Kong	0,1	0,1	6306,2	1986,9
Sydney	0,0	0,1	6255,2	1833,3
Singapour	0,2	0,1	6133,6	1285,1
Milan-Pavie	0,0	-0,1	5819,9	1966,5
Rome	0,0	0,0	5787,3	2079,4
Munich	-0,1	-0,1	5731,1	2167,9
Montreal	0,0	0,0	5696,4	1710,9
Melbourne	0,0	0,1	5695,8	1631,8
Baltimore	-0,1	-0,1	5588,4	2245,8
Barcelone	0,0	0,1	5525,6	1724,4
Total des 30 agglomérations	-1,9	-1,3	304773,9	95969,4
Total Monde	0,0	0,0	1101658,7	378098,3
Poids des 30 agglomérations en 2007* (%)			27,7	25,4

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: SCIEp (articles, recensions et lettres)

Tableau 20 – L'évolution du poids mondial des grands centres de production scientifique en matière de production et de co-production entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Ainsi, entre 2000 et 2007, la participation scientifique des 30 plus grands centres à la production de leur pays et à la production mondiale a été stable ou déclinante. Les grandes villes américaines, japonaises et allemandes, ainsi que Paris, Londres et Moscou ont perdu de leur importance relative. Alors qu'elle augmentait en valeur absolue

(+ 36,7 %), l'activité scientifique de ces centres n'a pas connu une croissance aussi rapide que celle intervenue par ailleurs (+ 50,3 %). C'est pourquoi, alors que 30 % des publications scientifiques mondiales était issu de ces 30 premières agglomérations en 2000, elles n'en produisent plus que 28 % en 2007.

Seules les grandes villes chinoises ont vu évoluer favorablement leur poids dans la production scientifique mondiale. En effet, le Tableau 20 montre que Beijing et Shanghai font partie des rares centres scientifiques (avec Séoul) dont la contribution relative à la production mondiale a progressé entre 2000 et 2007. Cette évolution n'est pas contradictoire avec le fait que ces agglomérations ont tendance à peser de moins en moins lourd dans la production scientifique de leur pays respectifs au cours des 30 dernières années. Tout cela traduit un double phénomène de déconcentration des activités des productions scientifiques :

- d'une part, les sites qui polarisaient autrefois l'activité scientifique de leur pays (généralement les « sites-capitales ») ont perdu un peu de leur monopole national suite aux politiques de décentralisation et au développement de nouveaux sites de recherche et d'enseignement supérieur ;

- d'autre part, de plus en plus de pays participent activement à l'entreprise scientifique mondiale abolissant le monopole des pays dits « développés » qui réalisaient traditionnellement la plus grande part de la production.

Si la Chine est le pays dont la participation croissante à la science mondiale a bousculé le plus visiblement la géographie de la production scientifique, de nombreux travaux montrent que ce n'est pas le seul. Au cours des années 2000, les analyses bibliométriques du WoS ont fait état d'un déclin de la place occupée par les pays d'Europe et d'Amérique du Nord (Adams & Pendlebury, 2010 ; UNESCO, 2010 ; Royal Society, 2011 ; Sexton, 2012) au profit de l'intégration à la « science mondiale » de pays dit « émergents » comme la Chine, l'Inde et le Brésil (Glänzel, Debackere, & Meyer, 2008 ; Huang, Chang, & Chen, 2012 ; Ponomariov & Toivanen, 2014). En considérant la partie la plus citée du WoS (Leydesdorff, Wagner, & Bornmann, 2014), en se limitant à la partie du WoS qui n'a pas été touchée par l'opération de « Regional Expansion », ou en ne considérant que le sous-ensemble des co-publications interurbaines (les activités collaboratives, soit 35 % des publications en 2007), il apparaît que ce phénomène de rééquilibrage est atténué mais ne disparaît pas, semblant confirmer un changement dans l'ordre établi. Le Tableau 21 résume ce phénomène. Ce tableau montre que parmi les 36 pays réalisant 95 % de la production scientifique mondiale en 2007, le poids des États-Unis, du Japon, de l'Allemagne, du Royaume-Uni, de la France et de la Russie a sensiblement diminué au profit de sept autres pays : la Chine (+ 5 points), la Corée du Sud (+ 1), le Brésil (+ 0,8), la Turquie (+ 0,8), l'Inde (+ 0,7), Taiwan (+ 0,6) et l'Iran (+ 0,6).

Pays les plus publiants en 2007*	Poids dans la production mondiale (%)		Production mondiale (ensemble des publications)		Co-production mondiale (ensemble des co-publications)	
	2000*	2007*	Evolution (%2000*-%2007*)	Nombre en 2007*	Evolution (%2000*-%2007*)	Nombre en 2007*
Etats-Unis	28,6	24,8	-3,8	272701,9	-3,0	101134,7
Chine	3,6	8,8	5,2	96474,5	4,1	27773,0
Japon	8,8	6,8	-2,1	74428,8	-1,9	24628,1
Allemagne	7,1	5,8	-1,4	63690,9	-1,1	25757,1
Royaume-Uni	7,4	5,6	-1,8	62139,7	-1,2	23189,7
France	5,2	4,2	-1,0	46632,3	-0,9	18870,6
Italie	3,5	3,6	0,1	39525,9	-0,2	15152,0
Canada	3,4	3,4	0,0	37447,5	0,1	12631,1
Inde	2,1	2,8	0,7	30487,9	0,6	6444,4
Espagne	2,4	2,7	0,2	29505,0	0,4	9699,3
Corée du Sud	1,6	2,6	1,0	28621,5	0,8	7891,9
Australie	2,1	2,2	0,0	23758,1	0,2	7520,1
Russie	2,9	2,0	-0,9	21833,6	-0,6	5200,6
Brésil	1,2	1,9	0,8	21231,5	0,7	7557,1
Taiwan	1,2	1,8	0,6	19827,5	0,6	5738,4
Pays-Bas	1,9	1,7	-0,2	18725,0	-0,1	8029,9
Turquie	0,7	1,5	0,8	16726,9	0,6	3968,7
Pologne	1,1	1,3	0,2	14512,5	0,0	3645,9
Suisse	1,3	1,2	-0,2	12967,9	-0,1	5367,8
Suède	1,5	1,1	-0,4	12642,6	-0,4	5418,6
Belgique	1,0	0,9	0,0	10110,4	0,0	4103,2
Israël	1,0	0,8	-0,2	8875,0	-0,2	2816,2
Iran	0,2	0,8	0,6	8787,1	0,4	2066,3
Grèce	0,5	0,8	0,2	8349,5	0,2	2374,1
Autriche	0,7	0,7	-0,1	7362,0	0,0	2546,9
Finlande	0,7	0,6	-0,1	6916,1	-0,1	2944,7
Danemark	0,7	0,6	-0,1	6743,4	-0,1	2906,4
Mexique	0,5	0,6	0,1	6379,7	0,0	2089,7
Singapour	0,4	0,6	0,2	6133,6	0,1	1285,1
République Tchèque	0,4	0,5	0,1	5882,7	0,1	1930,6
Portugal	0,3	0,5	0,2	5505,9	0,2	2104,6
Norvège	0,5	0,5	0,0	5248,7	0,0	2087,4
Argentine	0,4	0,4	0,0	4469,2	0,0	1446,2
Nouvelle-Zélande	0,4	0,4	-0,1	4122,5	0,0	1498,5
Hongrie	0,4	0,4	0,0	4024,0	-0,1	1426,9
Afrique du Sud	0,4	0,4	0,0	3943,3	0,0	1361,5
Total des 36 pays	96,0	95,0	-1,0	1046734,6	-0,8	360607,4
Monde	100,0	100,0	0,0	1101658,7	0,0	378098,2
Poids des 36 pays en 2007* (%)				95,0		95,4

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: SCIEp (articles, recensions et lettres)

Tableau 21 – L'évolution du poids mondial des grands pays de production scientifique en matière de production et de co-production entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Les cartes de l'évolution de la production mondiale par agglomération permet de visualiser ce phénomène (deux types de projections différents et de discrétisations sont utilisés pour témoigner du même évènement) (Figure 58).

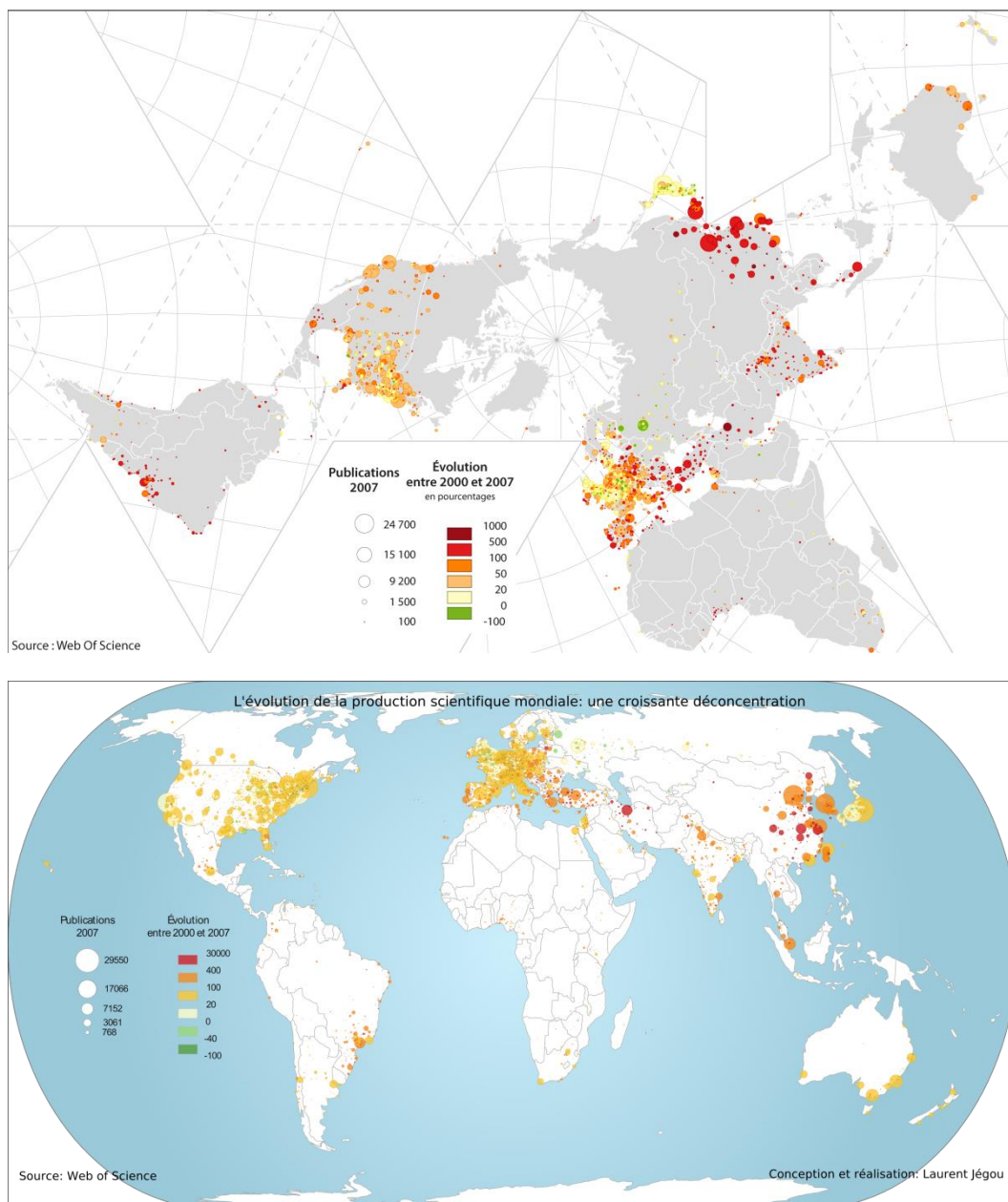


Figure 58 — Cartes de la production scientifique mondiale et son évolution entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Ces cartes confirment qu'il s'agit d'un double phénomène ou plus exactement d'un phénomène multiscalaire, c'est-à-dire que le rééquilibrage intervient à plusieurs échelles. En effet, lorsqu'on regarde uniquement ce qu'il se passe à l'intérieur de l'« Espace Européen de la Recherche », on constate que la moindre croissance des publications scientifiques au Royaume-Uni, en Allemagne et en France, a permis aux agglomérations d'Italie, d'Espagne et des Pays-Bas de contribuer davantage aux publications euro-

péennes ; et qu’au sein de ces différents pays les villes qui ont le plus participé à la croissance ne sont ni les capitales, ni les villes les plus publiantes mais les villes « secondaires » — puisque les cercles les plus foncés ne sont pas les plus grands (Figure 59).

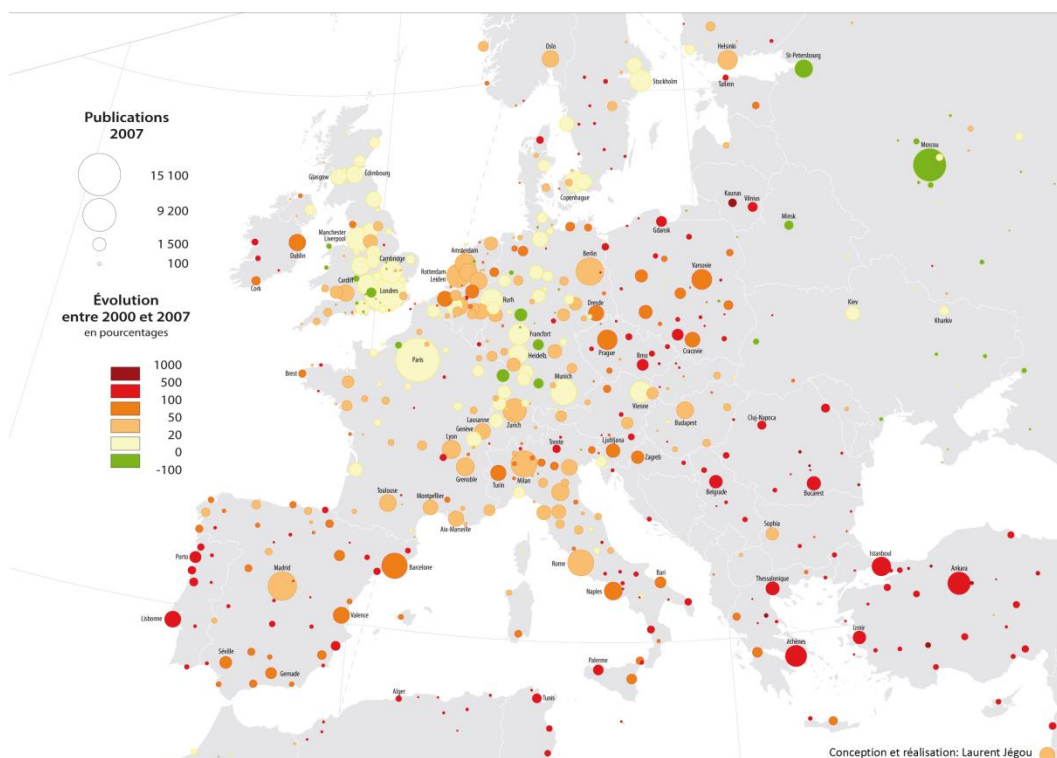


Figure 59 – Carte de la production scientifique européenne et son évolution entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

En plus de la déconcentration de la production scientifique à plusieurs niveaux, les tableaux précédents (Tableaux 20 et 21) témoignent d’une déconcentration des activités collaboratives, et plus précisément d’une répartition plus équitable des co-productions interurbaines. En effet, le Tableau 20 montre que, même si elle est légèrement atténuée, l’importance accrue de Beijing dans le total mondial est également nette lorsqu’on se limite aux co-productions interurbaines ; et le Tableau 21 que la part de la Chine augmente presque autant parmi les co-productions interurbaines qu’au sein de l’ensemble de la production. Cette corrélation entre le degré de concentration géographique de la production scientifique et celui des co-productions mérite d’être soulignée — la corrélation de Pearson entre l’évolution du poids dans la production et l’évolution du poids dans la co-production est significative au niveau des villes et au niveau des pays ($p\text{-value} = 2,2\text{e-}16$). Le coefficient de corrélation est égal à 0,98 au niveau des grandes villes et 0,99 au niveau des pays. Il est important de le noter car malgré la déconcentration mondiale de la pro-

duction, il aurait pu subsister un système d'échange très hiérarchisé comme le suppose la théorie répandue sur le renforcement du caractère nodal des « villes globales ». Dans un tel système, seuls quelques grands centres d'excellence ont la stature permettant de participer aux échanges mondiaux. Cette idée est souvent véhiculée dans les discours territorialisés sur la recherche, par exemple à travers l'affirmation que certains grands pôles ont une « vocation internationale ».

Le fait que la déconcentration mondiale de la production scientifique s'accompagne d'une déconcentration des co-productions entre agglomérations suggère que l'« agrandissement du monde » lié à l'intégration progressive de nouveaux espaces dans le système scientifique ne s'accompagne pas nécessairement d'un renforcement de la centralité des plus grandes villes, souvent présentées comme les ponts ou *hubs* indispensables à la cohésion de l'ensemble. La moindre progression ou la baisse relative de la centralité des plus grandes villes ne permet cependant pas de statuer sur le renforcement éventuel de leur rôle d'intermédiaires au niveau mondial. Le nombre brut ou relatif de collaborations par ville ne dit rien de la structure du réseau des collaborations et *a fortiori* de la nature des nouveaux liens qui se sont tissés entre les deux périodes. Ces résultats ne font qu'informer du fait qu'un processus de densification et de diffusion a eu lieu et que les plus grandes agglomérations scientifiques n'y ont pas joué un rôle majeur. Sans faire l'effort de décomposer les collaborations, il n'est donc pas possible de saisir toutes les facettes de la déconcentration de l'activité scientifique.

Afin de vérifier dans quelle mesure la globalisation des activités scientifiques (entendue comme l'augmentation du nombre d'agglomérations participant à la production mondiale) a pu porter atteinte au caractère structurant des cadres nationaux et, plus précisément, aboutir à l'internationalisation des activités scientifiques et l'effacement des frontières si souvent évoquées dans la littérature sur les sciences, il convient maintenant d'opérer une distinction entre les liens de collaborations intra-nationaux et les liens de collaborations internationaux.

1.3. Nationalisation et internationalisation des activités scientifiques

L'analyse spatiale des données bibliographiques en dynamique a mis en évidence un phénomène de rééquilibrage entre pays, au profit de plusieurs pays « émergents » (Chine, Corée du Sud, Brésil, Turquie, etc.), ainsi que, dans la plupart des pays, une déconcentration des publications au profit d'agglomérations scientifiques autrefois « secondaires » et donc une perte d'hégémonie des capitales et plus grandes villes depuis les années 1980. Pour étudier la géographie des activités scientifiques, non plus seulement à travers la répartition géographique de la production scientifique, mais en tenant compte de la struc-

ture géographique des réseaux de collaborations scientifiques, on considère la matrice des co-publications interurbaines extraite des données du SCI Expanded. Les résultats obtenus suggèrent que la « globalisation » des activités scientifiques (entendue cette fois comme l'augmentation du nombre de collaborations scientifiques) s'est accompagnée au cours des années 2000 d'un renforcement de la cohésion des espaces scientifiques nationaux de plusieurs pays. Il s'agit maintenant d'en rendre compte et de chercher à comprendre la diversité des cas de figure observés — en ayant à l'esprit que la croissance remarquable des publications intervenue entre 2000 et 2007 (+ 46 %) a été moins forte que celle des co-publications qui sont passées de 231 590 à 379 000 (+ 64 %). En effet, comme l'indique le Tableau 22, le nombre de publications signées depuis au moins deux villes a augmenté entre 2000 et 2007, en passant de 31 % à 35 % de l'ensemble des publications du SCI Expanded. Sans faire l'effort de décomposer les données de co-publications par ville et par pays, on constate qu'il n'est pas possible de déterminer si cette augmentation traduit plutôt une densification des échanges à l'intérieur ou entre les pays. Contrairement à ce qui est souvent affirmé, ce tableau montre qu'il n'y a pas eu de tendance globale à l'internationalisation des échanges scientifiques au cours de la dernière décennie. Les publications multi-localisées ont toutes augmenté qu'elles soient intra- ou internationales. Dès lors, il convient de décomposer les données de co-signatures pour caractériser la croissance des collaborations des villes les plus dynamiques en matière de production scientifique mais aussi des principaux pays.

Publications (articles, recensions et lettres) signées depuis:	% en 2000*	% en 2007*
Une ville	68,8	64,9
Plusieurs villes du même pays	15,3	17,3
Plusieurs villes, plusieurs pays	16,0	17,8
Total (%)	100,0	100,0
<i>Nombre de publications en valeur absolue</i>	<i>763203,0</i>	<i>1117566,3</i>

*Comptage complet non fractionné, moyenne mobile sur 3 ans. Source: SCI Exp

Tableau 22 — Nombre d'agglomérations et de pays par publication en 2000 et 2007.
Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

À l'encontre de l'idée d'un renforcement du caractère global des plus grands centres de production scientifique, le Tableau 23 indique que très peu d'entre eux ont consacré une part plus importante de leurs collaborations aux connexions internationales entre 2000 et 2007. En effet, seule une partie des grands centres a vu diminuer la part nationale des collaborations au profit des collaborations internationales entre 2000 et 2007 (Tokyo, New-York, Boston, Londres, Los-Angeles, Baltimore). Les grands centres chinois

ainsi que Munich, Rome, Moscou, Madrid, Barcelone, Montréal et Melbourne sont dans le cas inverse : on assiste à une progression de leurs collaborations intra-nationales au détriment de leurs collaborations internationales.

Agglomérations les plus publiantes en 2007*	Part des collaborations nationales (%)		Evolution (%2000*-%2007*)	Nombre de co-publications en 2007*
	2000*	2007*		
Tokyo	68,8	67,7	-1,1	6943,5
Beijing	57,4	65,7	8,3	5899,5
Paris	36,2	36,4	0,3	5705,5
New York	63,7	62,7	-1,0	5582,5
Boston	62,8	59,7	-3,1	5542,9
Londres	39,7	34,0	-5,6	5115,5
Washington-Bethesda	67,8	67,4	-0,4	4948,7
Baie de San-Francisco	63,9	63,7	-0,1	4536,0
Kyoto-Osaka	71,7	70,9	-0,8	4211,8
Los Angeles	64,0	62,5	-1,5	4060,2
Seoul	52,1	52,7	0,6	3269,4
Chicago	68,9	68,1	-0,8	2670,8
Shanghai	59,9	62,8	2,9	2639,8
Berlin	39,7	39,5	-0,2	2584,0
Philadelphie	70,0	69,4	-0,7	2554,7
Durham Research Triangle	71,9	71,3	-0,6	2515,2
Toronto	31,5	32,1	0,6	2358,6
Baltimore	71,1	69,4	-1,7	2245,8
Munich	38,7	41,5	2,8	2167,9
Taipei	52,8	64,1	11,2	2091,5
Rome	50,4	51,9	1,4	2079,4
Moscou	23,3	29,7	6,4	2077,5
Madrid	33,8	37,5	3,8	2029,6
Hong Kong	48,9	58,5	9,6	1986,9
Milan-Pavie	50,2	49,5	-0,7	1966,5
Sydney	34,3	33,7	-0,5	1833,3
Barcelone	30,9	33,0	2,2	1724,4
Montreal	31,3	32,5	1,2	1710,9
Melbourne	33,1	36,9	3,8	1631,8
Singapour	0,0	0,0	0,0	1285,1
Total des 30 agglomérations	57,6	56,5	-1,1	95969,4
Total Monde	53,7	55,0	1,4	378098,3
<i>Poids des 30 agglomérations en 2007* (%)</i>				25,4

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: *SCIExp (articles, recensions et lettres)*

Tableau 23 – L'évolution de la part nationale des collaborations des grands centres de production scientifique entre 2000 et 2007. Source : *SCI Expanded (articles, reviews, letters)*, moyenne mobile sur 3 ans

En fait, en rapportant ces observations à celles obtenues au niveau des pays (Tableau 24) on remarque que l'évolution du profil de collaboration des grandes villes va dans le même sens que l'évolution propre à leur pays d'appartenance ($p\text{-value} = 6,972\text{e-}13$; coefficient de corrélation = 0,92). Cela suggère que le fait de participer au « club » des plus grandes villes (en matière de production scientifique) n'empêche pas d'être sensible à des logiques nationales et qu'au fond, ce n'est pas la taille de la ville ou la taille du pays qui détermine son profil de collaboration.

Ainsi, les grandes villes qui se caractérisent par une franche tendance à la nationalisation (c'est-à-dire, qui ont augmenté la part de leurs collaborations avec les villes de leur propre pays) sont celles qui appartiennent aux pays les plus marqués par une nationalisation des activités collaboratives au cours de la dernière décennie, en l'occurrence : la Chine, Taiwan, l'Espagne, et la Russie. La grande ville dont le repli domestique a été le plus fort avec Taipei (+ 11 points) est Moscou dont la part des relations intra-nationales a augmenté de 6 points entre 2000 et 2007, au détriment de ses collaborations internationales. Contrairement à Taipei dont la moitié des relations étaient déjà intra-nationales en 2000, Moscou est la ville dont la part de relations intra-nationales était la plus faible en 2000 (23 %), sans doute parce que la production scientifique du pays était autrefois très centralisée. Cette évolution est notable mais doit être relativisée puisque la Russie a connu une augmentation sporadique de sa production scientifique entre 2000 et 2007 (+ 0,9 % soit seulement 200 publications supplémentaires), alors que son système scientifique était en pleine restructuration (Milard, 2008). Pour expliquer le repli domestique de Taipei, la forte densification intra-nationale de l'île dont témoigne la chute du poids relatif de Taipei dans la production nationale entre 2000 et 2007 (-7 points d'après le Tableau 19) est un argument intéressant, auquel s'ajoute la possibilité d'un isolement croissant dont serait responsable le fort développement de la Chine continentale. Par ailleurs, il apparaît que les métropoles américaines ont suivi plus ou moins franchement la tendance de leur pays à l'internationalisation : l'ouverture au monde de Boston a été beaucoup plus franche que celle de New-York, et plus encore que celle de Washington et San-Francisco. Mais, contrairement à ce qu'il s'est passé dans les autres pays, en particulier dans les pays les plus centralisés, la déconcentration de la production scientifique intervenue aux États-Unis entre 2000 et 2007 a peu affecté le poids respectif de chacune de ces grandes métropoles américaines.

Puisque dans certains cas comme celui de Taipei, le développement considérable des liens intra-nationaux coïncide avec une forte diminution du monopole de la ville-capitale dans la production nationale, on peut faire l'hypothèse qu'il existe une relation entre la perte de monopole de certains centres et la tendance à la « nationalisation » des collaborations de ces derniers ; ce qui expliquerait à l'inverse que les quelques centres ayant accru leur caractère polarisant à l'échelle nationale entre 2000 et 2007, comme Tokyo (+ 1,4 points dans la production nationale d'après le Tableau 19), se soient davantage internationalisés.

Pays ayant le plus de collaborations en 2007*	Part des collaborations nationales (%)		Evolution (%2000*-%2007*)	Nombre de co-publications en 2007*
	2000*	2007*		
Etats-Unis	69,5	68,6	-0,9	101134,7
Chine	64,2	70,7	6,6	27773,0
Allemagne	47,1	47,6	0,4	25757,1
Japon	74,4	72,9	-1,5	24628,1
Royaume-Uni	49,0	44,0	-5,0	23189,7
France	46,8	46,8	0,0	18870,6
Italie	56,3	56,8	0,5	15152,0
Canada	41,2	41,3	0,2	12631,1
Espagne	43,9	47,5	3,6	9699,3
Pays-Bas	44,5	46,3	1,9	8029,9
Corée-du-Sud	62,6	63,9	1,3	7891,9
Brésil	55,2	69,1	13,8	7557,1
Australie	38,5	39,3	0,8	7520,1
Inde	56,8	63,8	7,0	6444,4
Taiwan	66,3	75,1	8,8	5738,4
Suède	39,8	38,6	-1,2	5418,6
Suisse	22,2	22,5	0,3	5367,8
Russie	29,7	36,6	6,9	5200,6
Belgique	27,8	31,8	4,0	4103,2
Turquie	55,2	71,7	16,4	3968,7
Pologne	29,0	39,9	10,9	3645,9
Finlande	48,5	49,5	1,0	2944,7
Danemark	32,3	34,4	2,1	2906,4
Israël	31,9	37,6	5,7	2816,2
Autriche	18,8	20,0	1,2	2546,9
Grèce	33,6	45,2	11,6	2374,1
Portugal	28,5	38,9	10,3	2104,6
Mexique	36,1	43,2	7,1	2089,7
Norvège	35,2	35,6	0,4	2087,4
Iran	49,8	63,5	13,7	2066,3
République Tchèque	24,7	35,8	11,1	1930,6
Nouvelle-Zélande	37,0	31,7	-5,3	1498,5
Argentine	33,5	36,0	2,5	1446,2
Hongrie	22,8	30,3	7,5	1426,9
Afrique du Sud	37,1	33,0	-4,1	1361,5
Singapour	0,0	0,0	0,0	1285,1
Total des 36 pays	55,2	56,7	1,5	360607,4
Monde entier	53,7	55,0	1,4	378098,2

*Comptage fractionné (WNC), moyenne mobile d'ordre 3. Source: *SCIExp (articles, recensions et lettres)*

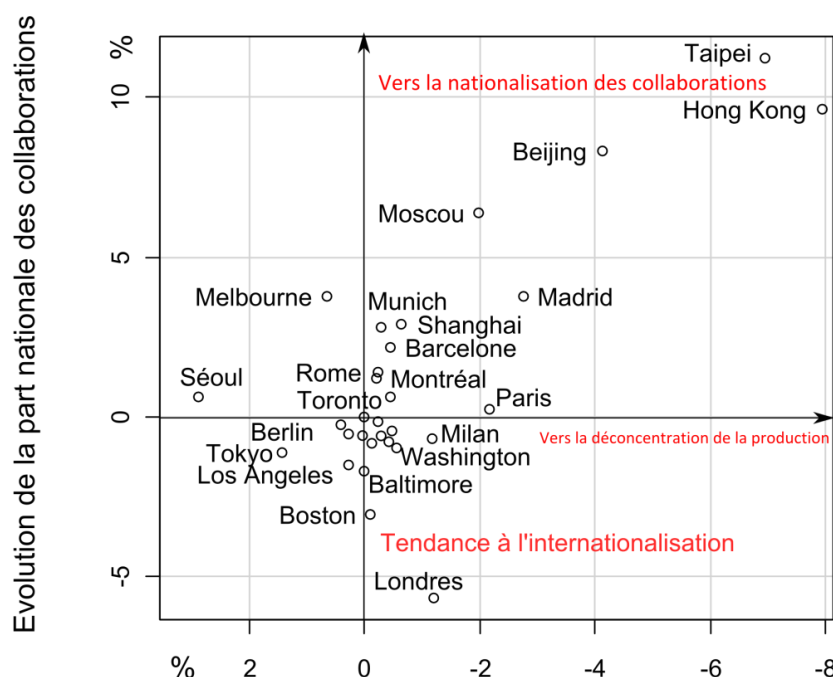
Tableau 24 – L'évolution de la part nationale des collaborations des grands pays de production scientifique entre 2000 et 2007. Source : *SCI Expanded (articles, reviews, letters)*, moyenne mobile sur 3 ans

En effet, la perte du monopole d'une ville dans la production de son pays suggère que des centres alternatifs ou secondaires ont renforcé leur activité de production scientifique. Ces espaces en développement se présentent comme autant de nouveaux collaborateurs potentiels pour la ville-centre. La déconcentration de la production scientifique d'un pays favoriserait le polycentrisme⁹⁷ de son système national de recherche.

⁹⁷ On fait référence à la notion de polycentrisme en y intégrant des implications en matière d'organisation spatiale et, en particulier, de fonctionnement réticulaire. Il s'agit donc de distinguer entre le

Cette hypothèse est confirmée par la Figure 60 qui montre l'existence d'une relation significative entre l'évolution de leur poids dans la production scientifique de leur pays et l'évolution de la part nationale de leurs collaborations. Plus la ville a perdu de poids dans la production nationale, plus elle a développé de collaborations interurbaines à l'intérieur de son espace national.

Corrélation entre l'évolution du poids dans la production nationale
et la part nationale des collaborations
(Les 30 agglomérations qui ont le plus publié en 2007* dans le SCI Expanded)



Evolution du poids dans la production nationale entre 2000* et 2007* (%)

*moyenne mobile sur 3 ans, whole normalized counting.
Source: SCI Expanded (articles, revues et lettres)

1. Régression linéaire: coefficient de corrélation multiple (r^2): 0,5628
p-value: 1,811e-06
2. Coefficient de corrélation de Pearson: -0,75
p-value = 1,811e-06

MM

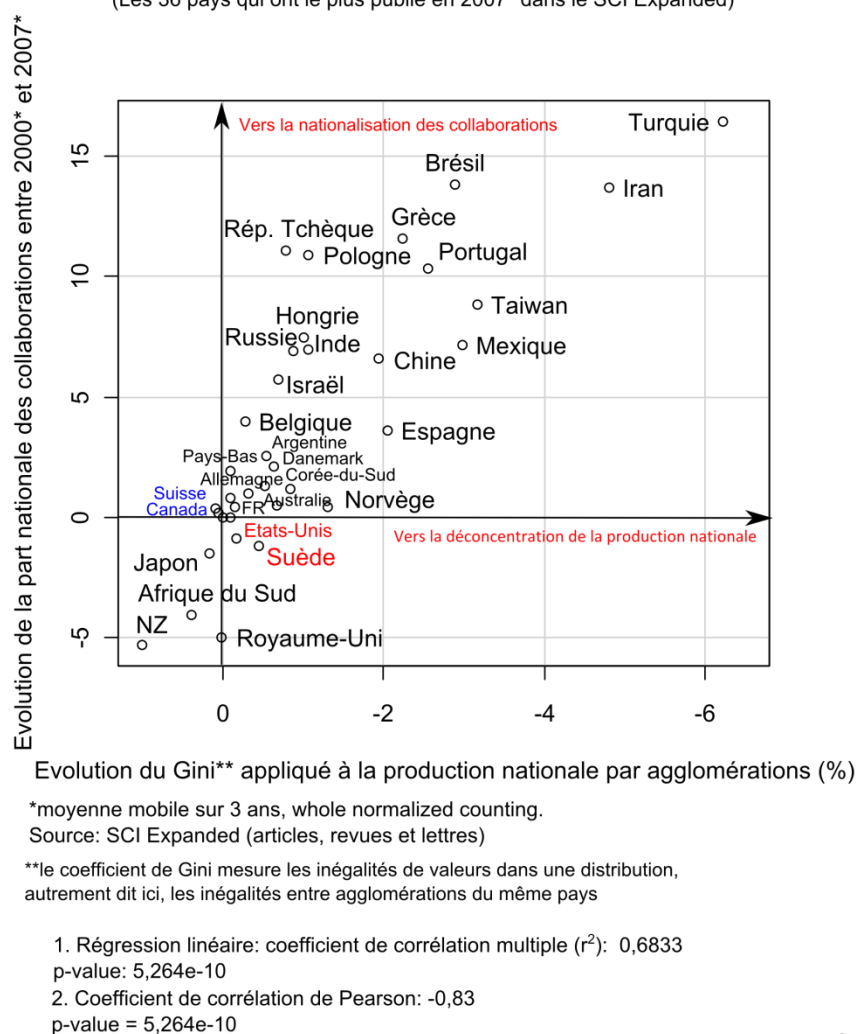
Figure 60 – Corrélation entre l'évolution du poids dans la production scientifique et la part nationale des collaborations des 30 premiers centres scientifiques mondiaux.

Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

fait que plusieurs pôles puissent co-exister dans le même pays, et le fait que ces pôles soient en relation et fassent système.

Si cette relation est valable pour les 30 premières agglomérations du monde, il convient de vérifier qu'elle se maintient en considérant l'ensemble des agglomérations des pays les plus publiants du monde. Pour cela, il suffit de tester la relation entre l'évolution du degré de concentration de la production scientifique interne à chaque pays et la tendance à la nationalisation ou l'internationalisation des échanges scientifiques propres à tous ces pays entre 2000 et 2007. Ce faisant, on trouve que l'évolution de la valeur de l'indice de concentration de la production scientifique (indice de Gini) des 36 pays les plus publiants du monde parvient significativement à expliquer l'évolution du penchant des agglomérations de ces pays à accroître leurs échanges intra ou internationaux (Figure 61).

Corrélation entre l'évolution du Gini** appliqué à la production nationale et la part nationale des collaborations
(Les 36 pays qui ont le plus publié en 2007* dans le SCI Expanded)



MM

Figure 61 – Corrélation entre l'évolution du degré de concentration de la production nationale par agglomération (indice de Gini) et la part nationale des collaborations des 36 premiers pays du monde. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Bien que ce résultat soit significatif, il convient d'apporter quelques bémols. La relation ainsi établie entre déconcentration de la production scientifique et nationalisation des collaborations interurbaines ne permet, en effet, d'expliquer que 68 % de la tendance des agglomérations à collaborer davantage à l'intérieur ou à l'extérieur de leurs cadres nationaux.

Plusieurs nuances méritent donc d'être apportées pour éviter les conclusions hâtives et compléter cette analyse.

Premièrement, cette explication porte sur les changements intervenus entre 2000 et 2007 dans la répartition de la production scientifique et dans le réseau mondial des collaborations scientifiques. Or, dans beaucoup de pays, ces changements ont été minimes, comme en témoigne la concentration des points autour de l'origine (Figure 61). L'indice de concentration comme la part nationale des collaborations ont été très stables dans plus de la moitié des pays. En particulier, les pays les plus anciennement impliqués dans l'activité scientifique comme la France, l'Allemagne, les pays scandinaves, le Canada, l'Australie, la Suisse et l'Italie ont connu des évolutions très modérées. L'évolution de leur profil de collaboration et l'évolution du degré de concentration de leur système national de recherche ont été caractérisées par une forte inertie, surtout au regard de ce qu'il s'est passé en Chine, en Turquie, au Brésil pendant cette courte période. Les « vieux » pays ont donc bien connu une croissance de leur activité scientifique mais celle-ci n'est pas venue bouleverser l'organisation de leur système de recherche.

Deuxièmement, on ne peut qu'être frappé par le fait que les rares pays à avoir manifesté une tendance à l'internationalisation de leurs activités collaboratives entre 2000 et 2007 sont, à l'exception du Japon et de la Suède, des pays de tradition anglophone : Afrique du Sud, Royaume-Uni, Nouvelle-Zélande et, dans une moindre mesure, États-Unis. Ces pays enregistrent un comportement proche en dépit de leur distance géographique et de la diversité de leur système national de recherche (différentes tailles, différents degrés de centralisation). Or, comme ces pays partagent une histoire commune, cela suggère que les facteurs historiques et linguistiques méritent d'être considérés pour comprendre l'évolution de la géographie des activités scientifiques. Au-delà, pour expliquer la tendance de ces pays à s'ouvrir davantage aux collaborations internationales, il importe d'avoir à l'esprit le biais linguistique de la base de données qui a tendance à visibiliser la production domestique de ces différents pays au détriment de leur production internationale (au contraire, dans les pays non-anglophones comme la France, les publications ont d'autant plus de chances d'être en anglais et donc recensées dans le SCI, qu'elles sont le fruit d'une collaboration internationale). Ainsi, les États-Unis, très grand pays à la fois en nombre de lieux de science et de publications scientifiques, étaient apparemment caractérisés par un comportement assez autarcique en matière de production scientifique au cours

des années 1950-1970, qui a commencé à s'atténuer dans les années 1980 (Frame & Narin, 1988). Ce mouvement d'ouverture des scientifiques américains semble être toujours en marche, mais aux explications proposées par Frame et Narin qui regardaient les évolutions du système scientifique mondial, il faut ajouter les effets propres à la source mobilisée. D'abord très centrée sur les productions américaines du fait de son origine, cette source a petit à petit élargi son périmètre pour intégrer davantage de revues étrangères. Parallèlement, de plus en plus de revues étrangères se sont mises à publier des articles en anglais, se donnant ainsi les moyens d'être recensées, ce qui peut également avoir pour effet de « visibiliser » les productions domestiques des pays non-anglophones dont seules les co-productions internationales étaient autrefois visibles dans le SCI.

Troisièmement, contrairement à ce que pourrait laisser penser l'établissement d'un lien entre la déconcentration de la production scientifique intervenue dans plusieurs pays et le renforcement de leur cohésion nationale à travers l'augmentation des collaborations nationales interurbaines, il n'existe pas de relation entre la performance des pays en matière de production scientifique et leur degré de concentration. En effet, le lien que nous avons mis en évidence n'est que la tendance mécanique de la géographie des collaborations scientifiques à suivre l'évolution de la géographie de la production. Lorsque la géographie de la production se déconcentre dans un espace donné, des collaborations se développent avec les sites ayant bénéficié de cette déconcentration. Tout comme la concentration des recherches ne donne pas lieu à un surplus de production (puisque l'effet de masse critique n'est qu'une illusion), la déconcentration ne favorise pas l'efficacité des chercheurs. L'idée d'une relation entre le polycentrisme d'un système de recherche et son efficacité était chère à Joseph Ben-David, mais n'a pas été démontrée (Ben-David, 1991). Pour Ben-David, le polycentrisme devait expliquer la meilleure performance du système américain sur le système soviétique, ce qui était surtout motivé par des impératifs idéologiques. À l'encontre de cette croyance, le Tableau 25 montre que le niveau de production des pays est indépendant de leur degré de concentration : le Royaume-Uni et l'Allemagne produisent le même nombre de publications alors que le système allemand est beaucoup moins concentré que le système britannique. De même, alors que l'Australie et la Russie ont une hiérarchie urbaine comparable, ces pays ont connu des dynamiques très différentes entre 2000 et 2007.

En fait, si on parvenait à établir un lien entre le dynamisme scientifique (croissance de la production) et la densification des systèmes nationaux de recherche, il s'expliquerait par le dynamisme des pays en forte croissance scientifique ayant traversé des reconfigurations importantes au cours des décennies précédentes (Chine, Brésil, Inde, Taiwan, Pologne, Turquie, Grèce, République Tchèque, Iran). Ces pays ont été marqués par des politiques de développement de l'activité scientifique et, plus largement, socio-économique

à plusieurs niveaux et les chercheurs de ces pays ont été incités à occuper une place plus grande dans le système mondial. La nette progression des liens intra-nationaux dans ces pays au détriment de leurs échanges internationaux entre 2000 et 2007 traduit une progressive autonomisation de leur système de recherche lié aux efforts qui ont été faits au niveau de leurs États pour disposer d'une « science nationale ».

Nombre de publications en 2007*			50%	80%	95%
Etats-Unis	272702	Nombre de villes requis pour atteindre les seuils de 50, 80 et 95% du nombre d'articles dans le pays considéré	19	70	162
Chine	96475		6	16	36
Royaume-Uni	62140		6	16	30
Japon	74429		2	9	31
Allemagne	63691		10	24	41
France	46632		4	14	34
Canada	37448		4	11	21
Australie	23758		2	5	11
Brésil	21231		5	18	47
Russie	21834		2	11	36
Inde	30488		8	36	92
Espagne	29505		4	17	33
Suisse	12968		2	4	6
Suède	12643		3	5	14
Iran	8787		1	7	22
Mexique	6380		1	10	28
Nouvelle-Zélande	4122		2	4	6
Afrique-du-Sud	3943		2	4	11

*Comptage fractionné (WNC), moyenne mobile sur 3 ans.

Source: SCIEp (articles, recensions et lettres)

Tableau 25 – La concentration métropolitaine de la science dans les 11 premières nations scientifiques. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Dans l'ensemble, ces premiers résultats montrent que la tendance à la « nationalisation » des activités scientifiques n'entre pas en contradiction avec leur tendance à « l'internationalisation » puisque l'ensemble des collaborations scientifiques se développe en même temps que le système scientifique s'étoffe au profit d'espaces nationaux et locaux autrefois peu visibles et moins impliqués dans la recherche. L'ensemble de la production augmente et, au sein de cet ensemble, la part des publications écrites en collaboration augmente d'autant plus. Cette croissance de l'activité scientifique bénéficie à certains es-

paces davantage qu'à d'autres et se manifeste différemment selon les espaces, justifiant la poursuite d'une approche multi-niveau. Après avoir insisté sur le maintien du caractère structurant du niveau national, il importe de considérer l'effet d'autres niveaux d'organisation pour saisir l'évolution de la géographie des collaborations scientifiques. En effet, si la croissance des relations interurbaines a accompagné la croissance des relations internationales dans la majorité des pays entre 2000 et 2007, sans que les plus grandes villes ne soient épargnées, il reste à déterminer comment ces changements ont impacté la structure d'ensemble et la hiérarchie du réseau mondial et, au-delà, à s'intéresser à la répartition des collaborations au sein de sous-systèmes pour mieux comprendre cette structure. Plus précisément, il s'agit d'une part, d'observer la densification du réseau de villes et ses effets à différents niveaux hiérarchiques ; et d'autre part de saisir la structure d'ensemble du système scientifique mondial, c'est-à-dire de détecter les niveaux d'organisation infranationaux et supranationaux au sein desquels se déploient préférentiellement des coopérations interurbaines.

2. Géographie du système mondial des collaborations scientifiques

« Le cloisonnement du monde tient plus aux barrières qui sont dans les esprits et aux orbites de rayonnement des carrefours qu'à toutes les caractéristiques physiques inscrites dans l'espace. » (Gottmann, 1951, p. 223)

2.1. La densification du réseau de villes

La déconcentration géographique des activités de production et de collaboration scientifiques peut s'expliquer par l'intégration progressive de nouveaux espaces (pays et villes) dans le système scientifique mondial. Effectivement, il s'agit sans doute d'un effet mécanique lié à « l'agrandissement du monde », mais c'est aussi, comme le suggèrent certains résultats de la section précédente, le résultat de la croissance plus rapide de l'activité enregistrée dans les espaces anciennement secondaires. Pour en témoigner, il est pertinent, contrairement à ce qui a été fait précédemment, de nous centrer sur un ensemble fixe d'agglomérations pour observer l'évolution de leur position dans le système de production scientifique qu'elles dessinent. Cela présente notamment l'avantage de se prémunir des artefacts auxquels peuvent contribuer les toutes petites unités d'analyse (les villes à une seule publication).

	Nombre (x)		Poids relatif		Taux de croissance ($x_{2007^*} - x_{2000^*} / x_{2000^*}$)
	2000*	2007*	2000*	2007*	
Production des 500 agglomérations les plus publiantes en 2007*	671965,9	960507,6	89,21	87,19	42,94
Production des villes non publiantes en 2000* (nombre de villes: 2425)	0,0	1736,6	0,0	0,16	--
Production des villes restantes publiantes en 2000*	81255,4	139414,4	10,79	12,65	71,58
Production totale	753221,3	1101658,7	100	100	46,26
Nombre de villes publiantes	7828	9321			19,07

*Comptage fractionné (WNC), moyenne mobile sur 3 ans. Source: SCIEmp (articles, recensions et lettres)

Tableau 26 – Évolution de la production des 500 agglomérations les plus publiantes en 2007 par rapport à l'évolution de la production totale entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

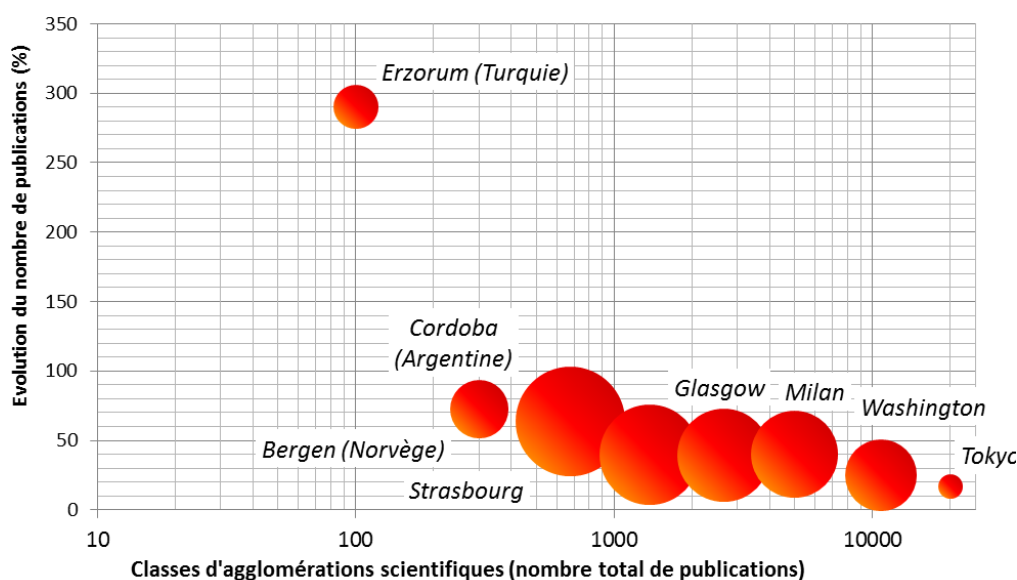
Ainsi, sans même inclure les nouveaux espaces de production à l'analyse – c'est-à-dire les 2425 villes publiantes en 2007 qui ne publiaient pas encore en 2000 (soit 0,16 % de la production en 2007) – et en se limitant à considérer la production des 500 villes les plus publiantes en 2007 (soit 87 % de la production en 2007, comme en témoigne le Tableau 26), on constate une déconcentration de la production scientifique interne à cet ensemble d'agglomérations qui s'accompagne d'une déconcentration de leurs collaborations (Tableau 27).

Agglomérations les 500 plus publiantes en 2007*	Part dans le total des publications (%)		Part dans le total des collaborations (%)		Evol.
	2000*	2007*	2000*	2007*	
Les 30 premières	33,82	31,73	33,15	31,31	↘↘
Les 100 premières	59,11	57,48	59,42	57,38	↘↘
Les 200 premières	78,50	77,10	78,93	77,21	↘↘
Les 300 premières	89,54	88,54	89,84	88,76	↘↘
Les 400 premières	96,41	95,42	96,77	95,91	↘↘
Les 500 premières	100	100	100	100	
Nombre total de publications et de copublications des 500 agglomérations	671965,937	960507,6137	178886,1034	277658,1514	

*Comptage fractionné (WNC), moyenne mobile d'ordre 3.
Source: SCIEmp (articles, recensions et lettres)

Tableau 27 – L'évolution de la répartition des publications et des co-publications entre les 500 agglomérations les plus publiantes du monde en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Même au sein de ce sous-ensemble de 500 agglomérations, qui a perdu un peu d'importance dans la production globale, il y a eu un effet de rééquilibrage ayant visiblement profité aux 100 dernières en termes de publications et de collaborations (+ 1 point). Cela dit, le choix de considérer des classes de villes de taille égale (100 villes par classe) ne permet qu'imparfaitement de rendre compte de la progression intervenue entre les deux périodes dans la mesure où la répartition des publications scientifiques suit une distribution de Lotka (Encadré 2). Pour mieux saisir l'évolution de la participation des villes en fonction de leur place dans cette hiérarchie bien particulière, il est préférable de procéder à une discrétisation par progression arithmétique. Elle est adaptée aux distributions qui, comme les distributions de publications scientifiques, comptent beaucoup de valeurs faibles (beaucoup de villes produisant peu de publications) et peu de valeurs fortes (peu de villes produisant beaucoup de publications).



*Les agglomérations indiquées sont représentatives de chacune des classes indiquées

Les cercles indiquent le nombre de publications supplémentaires entre 2000 et 2007 de chacune des classes d'agglomérations scientifiques. Ce sont les classes d'agglomérations scientifiques entre 400 et 6500 publications qui ont le plus contribué à la croissance générale des publications. En revanche, les classes d'agglomérations extrêmes - 100 et 20 000 - ont peu contribué à cette croissance.

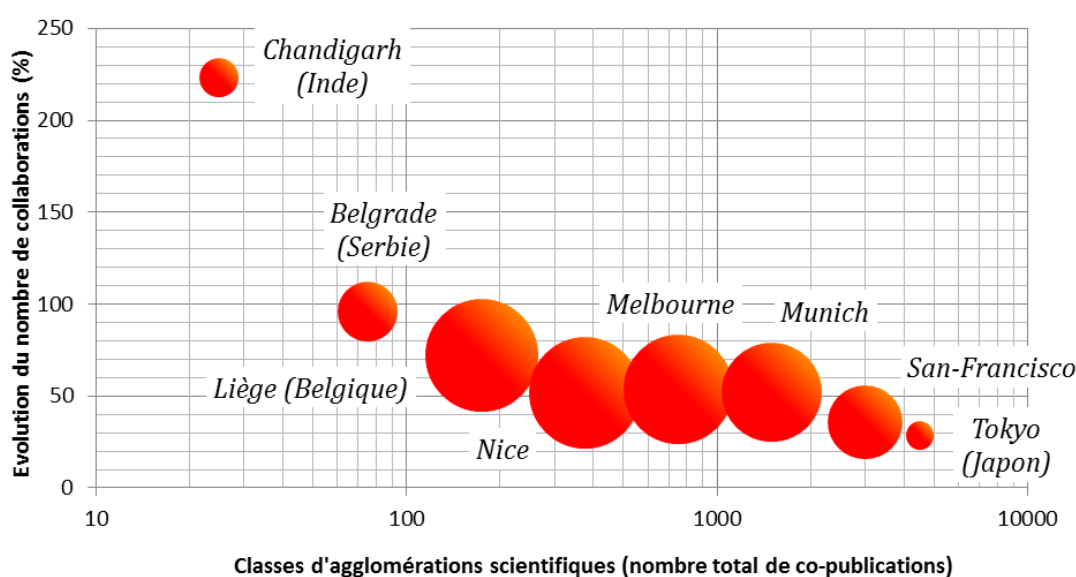
Source : Web of Science, SCIEp (articles, recensions et lettres), 1999-2001 et 2006-2008.

Conception: L. Jégou, M. Baron et M. Maisonobe. Réalisation:MM

Figure 62 – Quelles sont parmi les 500 premières agglomérations scientifiques celles qui participent à la croissance de la production des publications entre 2000 et 2007 ?
Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Ainsi, la Figure 62 montre que les agglomérations qui publiaient déjà entre 400 et 6500 publications en 2000 sont celles qui ont le plus contribué à la croissance générale de la production entre 2000 et 2007. Au contraire, bien que leurs taux de croissance ait été

très élevés, les villes ayant seulement produit moins de 400 publications en 2000 (les 2 dernières classes dont les représentants sont Erzurum et Cordoba) n'ont contribué qu'à 10 % du surplus de production réalisé entre les 2 périodes (la taille des cercles est proportionnelle à cette valeur), c'est-à-dire qu'elles ont contribué à la croissance dans les mêmes proportions que l'ont fait les villes ayant produit plus de 6500 publications en 2000 (les 2 premières classes dont les représentants sont Washington et Tokyo). Si la dispersion des valeurs et leur volume sont moins importants dans le cas des collaborations, on observe un phénomène analogue sur la Figure 63, mais avec des taux de croissance plus élevés, surtout parmi les classes centrales (entre 100 et 2000 collaborations en 2000).



*Les agglomérations indiquées sont représentatives de chacune des classes indiquées

Les cercles indiquent le nombre de publications supplémentaires entre 2000 et 2007 de chacune des classes d'agglomérations scientifiques. Ce sont les classes d'agglomérations scientifiques entre 100 et 2000 qui ont le plus contribué à la croissance générale des co-publications. En revanche, les classes d'agglomérations extrêmes - 25 et 4500 - ont peu contribué à cette croissance.

Source : Web of Science, SCIExp (articles, recensions et lettres), 1999-2001 et 2006-2008.

Conception: L. Jégou, M. Baron et M. Maisonobe. Réalisation: MM

Figure 63 – Quelles sont parmi les 500 premières agglomérations scientifiques celles qui participent à la croissance des collaborations entre 2000 et 2007 ? Source : SCIE Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Dans l'ensemble, le phénomène de croissance des co-publications a moins touché les « grandes » agglomérations que les « moyennes » agglomérations du système scientifique mondial en 2000. Comme pour la production scientifique, ce sont les entités du milieu de la hiérarchie en 2000 qui ont le plus participé à la croissance des co-productions

entre 2000 et 2007 (Figure 63). Les agglomérations scientifiques dont le nombre de collaborations était compris entre 100 (comme Liège) et 1000 (comme Melbourne) en 2000 ont réalisé à elles seules près de 65 % de la croissance totale entre 2000 et 2007.

Davantage, donc, que l'effet de l'intégration de nouveaux espaces de production entre les deux périodes étudiées, nous avons affaire à un phénomène de rééquilibrage. De nouveaux lieux de travail scientifique ont émergé, notamment pendant les années 1950-1960, qui alors qu'ils étaient secondaires pendant les années 1970-1990 participent de plus en plus nettement à la production scientifique de leur propre pays et, par la même occasion, du monde entier. Ce faisant, ils s'intègrent au réseau mondial des collaborations scientifiques dont le degré de concentration est de moins en moins élevé.

En plus de la croissance des connexions de chaque ville, la connectivité globale du réseau obtenue à l'aide des indices de densité et de degré (le premier mesure le nombre de connexions effectuées sur le nombre de connexions possibles et le second, le nombre moyen de connexions des villes) a augmenté de 25 % entre les deux périodes (Tableau 28). Ainsi, les trois quarts des connexions possibles entre les 500 agglomérations ont été réalisés au cours de la deuxième période. Il faut préciser que ces indicateurs ne tiennent pas compte de la valeur des liens (nombre de co-signatures entre deux villes) mais seulement de leur existence (y-a-t-il eu au moins une publication co-signée entre ces deux villes ?). En 2007, nous ne sommes donc plus très loin d'une situation où toutes les 500 villes considérées sont reliées entre elles par au moins une connexion. Cependant, le volume de collaborations varie d'un couple d'agglomérations à l'autre.

Les 500 agglomérations les plus publiantes en 2007*	2000*	2007*	Taux d'évolution (%)
Nombre de liens	74 567	92 519	24,1
Valeur des liens	178 886	277 658	55,2
Densité	0,6	0,75	25,0
Degré moyen	298,866	372,144	24,5

*Moyenne mobile sur 3 ans. Source: SCI Expanded (articles, recensions et lettres)

Tableau 28 – Évolution des mesures globales décrivant la connectivité du réseau des 500 agglomérations les plus publiantes du monde en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

D'après le Tableau 29, la participation des petites et moyennes agglomérations (moins de 500 articles co-produits en 2000 et moins de 750 en 2007) à la densification du réseau des 500 premières agglomérations n'a pas profité aux plus « grandes » villes du monde en nombre de collaborations. La part des co-productions scientifiques entre agglomérations principales et secondaires est restée globalement stable (elle représente envi-

ron la moitié des collaborations) ce qui achève de contredire l'idée d'un rôle de plus en plus central des grands pôles scientifiques (Tableau 29). De plus, on peut suspecter un mouvement de rééquilibrage de la structure hiérarchique des collaborations puisque la part des liens entre « moyennes et petites » villes progresse très légèrement (+1 point) au détriment de la part des liens entre « grandes » villes (-1 points). Tout se passe comme si la déconcentration géographique de la production et des collaborations scientifiques pouvait générer un tassement global de l'effet lié à la hiérarchie des agglomérations scientifiques dans la structuration du réseau de collaborations, mais la timidité de cette évolution suggère qu'il existe des résistances à de tels changements structurels. Cette évolution est, toutefois, plus nette lorsqu'on se limite au réseau des 202 agglomérations de l'Union Européenne à 28. Dans ce réseau, la part des liens entre « petites » villes est passée de 22 % à 25 % rejoignant presque la part des liens entre « grandes » villes (25 %). Cela fait écho à certaines études plus spécifiques portant sur le développement de la recherche dans les villes petites et moyennes et à l'intensification de leurs échanges (Milard, 2012 ; Levy *et al.*, 2013) (voir Partie 1.Chapitre 3.2).

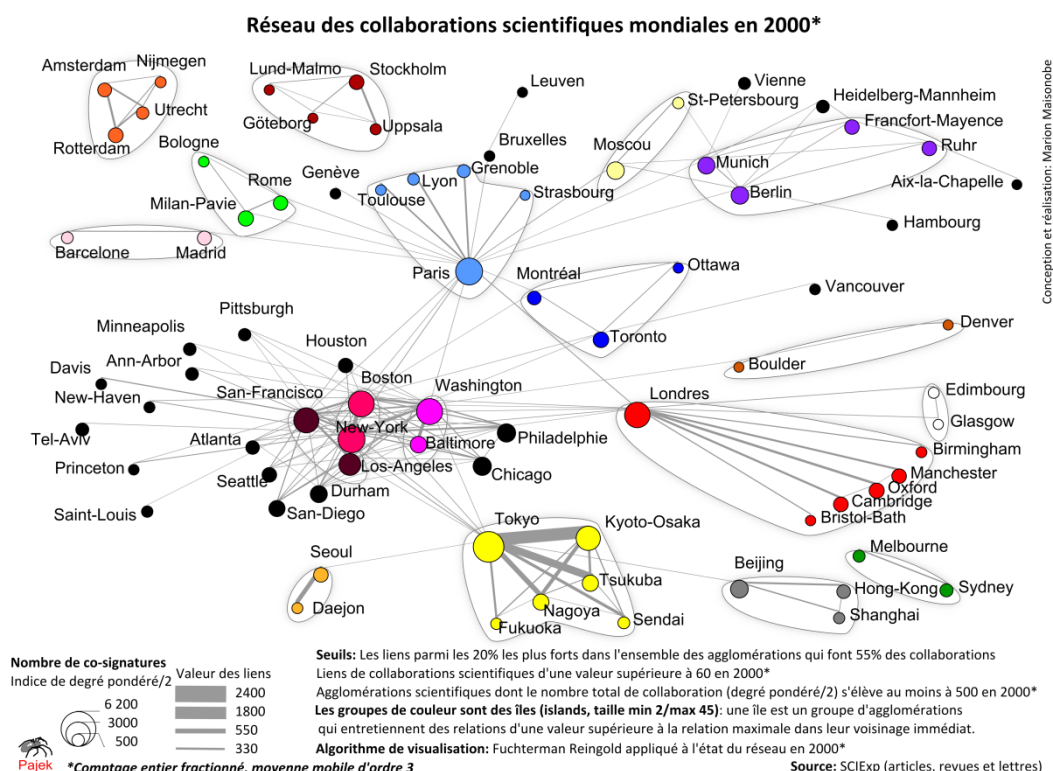
Parmi et entre les 500 agglomérations les plus publiantes du monde en 2007* (960 508 publications en 2007*)	% de collaborations scientifiques			Taux d'évolution (x 2007* - x 2000*)/x 2000*
	En 2000*	En 2007*		
Liens entre grandes villes**	33,12	32,01	↘	50,0
Liens entre petites et grandes villes**	47,14	47,06		55,0
Liens entre petites villes**	19,74	20,93	↗	64,5
Total (%)	100,00	100,00		
<i>Nombre de collaborations</i>	<i>178886,10</i>	<i>277658,15</i>		55,2

*Comptage fractionné (WNC), moyenne mobile d'ordre 3. Source: SCI Exp (articles, recensions et lettres)

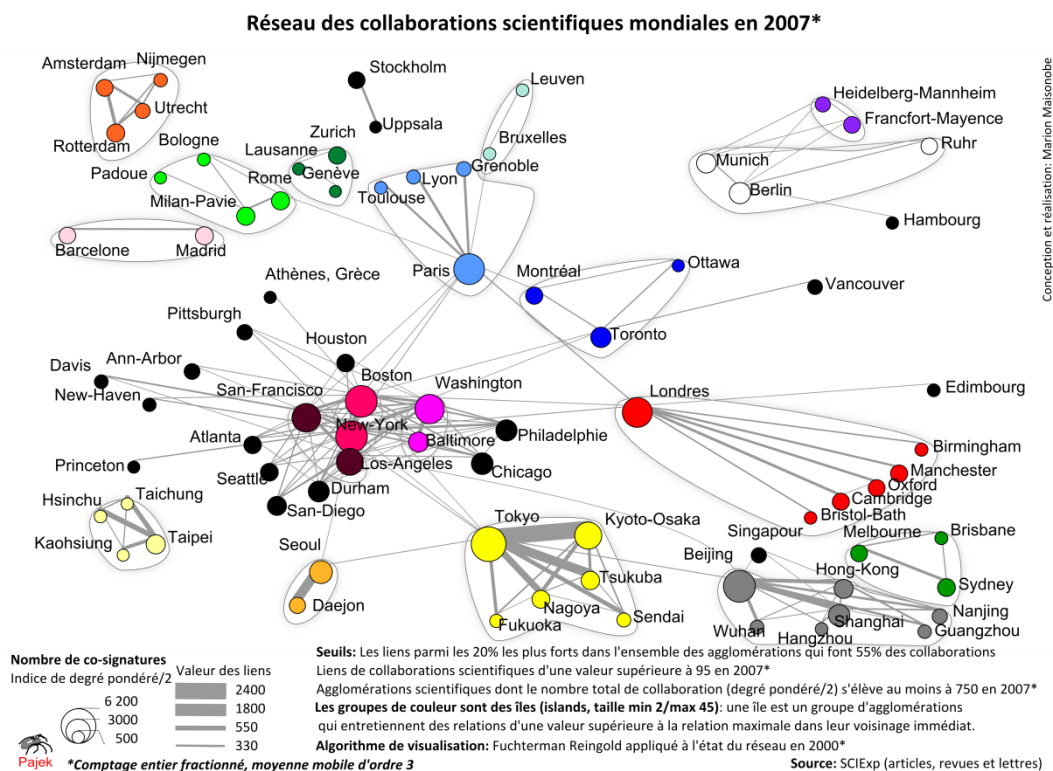
**410 "petites" villes ayant participé à moins de 500 co-publications en 2000* et 90 "grandes" villes, à plus de 500 co-publications. 406 "petites" villes ayant participé à moins de 750 co-publications en 2007 et 94 "grandes" villes à plus de 750

Tableau 29 – Évolution de la structuration des collaborations scientifiques parmi les 500 premières agglomérations en fonction de leur taille en nombre de collaborations.
Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Tandis que la structure hiérarchique apparaît globalement stable, la position des agglomérations dans cette hiérarchie est susceptible de se modifier ainsi que la répartition de leurs relations induisant des bouleversements au moins locaux ou régionaux dans la composition du réseau, comme en témoigne d'ailleurs le cas des agglomérations européennes.



Figures 64 et 65 – Le haut de la hiérarchie du réseau des 500 premières agglomérations en 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans



Ainsi, lorsqu'on se limite aux villes du haut de la hiérarchie et au réseau obtenu en ne conservant que les relations les plus fortes entre ces villes en 2000 puis en 2007, nous obtenons les deux représentations suivantes qui présentent plusieurs dissimilarités (Figures 64 et 65). Pour chaque année les agglomérations et les liens représentés obéissent au même critère (les villes retenues sont celles qui, lorsqu'elles sont ordonnées dans l'ordre décroissant, réalisent 55 % de l'ensemble des collaborations du réseau interurbain, et les liens représentés sont ceux qui, lorsqu'ils sont ordonnés dans l'ordre décroissant, représentent 20 % de l'ensemble des collaborations). Utiliser un critère identique pour les deux représentations garantit leur comparabilité. Ainsi, le fait que le nombre de villes retenues (environ 80) reste le même à une ville près en 2000 et en 2007 indique que le haut de la hiérarchie a conservé sa place dans le réseau mondial, c'est-à-dire qu'il faut environ le même nombre de villes pour réaliser la même part de collaborations aux deux dates (*nota bene* : les isolats ne sont pas représentés, c'est-à-dire les villes qui font partie des plus connectées, mais n'ont pas une seule relation dont la valeur dépasse le seuil retenu).

Toutefois, il importe de remarquer, pour contrer d'office l'hypothèse d'une stagnation entre les deux périodes, que les villes et les relations qui remplissent le critère retenu ne sont pas les mêmes. Entre 2000 et 2007, les villes occupant le haut de la hiérarchie et leurs connexions ont partiellement changé. Le changement le plus notable entre 2000 et 2007 est celui de l'intégration d'un grand nombre de villes chinoises en haut de la hiérarchie au détriment des villes suédoises, russes, et des villes du Colorado (Denver et Boulder). On observe aussi, à travers la variation de la taille des cercles, qui dépend du nombre total de collaborations de la ville (avec le reste des 499 villes), que ce sont les villes chinoises (Beijing, Shanghai, Hong Kong...) et plus généralement asiatiques qui ont connu la plus forte croissance globale de collaborations scientifiques. La déconcentration mondiale de la production scientifique au profit de la Chine et de la Corée du Sud a ainsi eu pour effet de bousculer le statut d'un certain nombre de lieux anciennement centraux.

Aussi, ces représentations montrent que la tendance à la « nationalisation » identifiée auparavant en tenant compte de toutes les relations interurbaines par pays se confirme en se limitant aux relations interurbaines les plus fortes. En effet, les liens les plus forts en 2007 sont plus souvent intra-nationaux que les liens les plus forts en 2000. Cela se vérifie en considérant le nombre de composantes connexes (ensembles de villes disjoints), qui passe de 4 en 2000 (la composante principale + les Pays-Bas, la Suède et l'Australie) à 7 en 2007 (la composante principale + les Pays-Bas, la Suède, l'Australie, la Suisse, l'Espagne et Taiwan). Dans l'ensemble, on remarque que les structures nationales se dégagent très bien des deux représentations. Il y a seulement quelques liens forts entre les plus grandes villes du monde qui contribuent à la cohésion de l'ensemble. Cette observa-

tion pourrait nous inciter à souscrire aux analyses sur les villes mondiales ; sauf qu'à l'encontre de la plupart des chercheurs de ce courant, nous récusons l'idée d'un mouvement de concentration de l'activité dans ces villes dictée par la globalisation.

De fait, nos résultats suggèrent que l'activité est de mieux en mieux répartie à l'échelle mondiale mais aussi nationale. La représentation graphique permet d'ailleurs de vérifier que les liens entre Boston, Paris, New-York et Londres sont moins forts et ont généralement évolué moins vite que ceux que ces villes entretiennent avec leur base domestique. Grâce à la méthode de comptage adoptée pour estimer le volume des collaborations, il n'est même pas nécessaire de contrôler par l'effet de taille pour observer que la préférence de ces « villes globales » (l'expression est de Saskia Sassen) va vers leurs partenaires domestiques (Tableau 30).

Les 15 premiers partenaires scientifiques des "villes globales" en volume de collaborations, 2007*

Tokyo	New-York	Paris	Londres	Beijing
Kyoto-Osaka	Boston	Lyon	Cambridge	Shanghai
Tsukuba	Washington	Grenoble	Oxford-Didcot	Hong-Kong
Nagoya	San-Francisco	Toulouse	Manchester	Changchun
Sendai	Philadelphie	Marseille-Aix	Bristol-Bath	Wuhan
Fukuoka	Los-Angeles	Montpellier	Birmingham	Lanzhou
Sapporo	Chicago	Bordeaux	<u>Paris</u>	Nanjing
Shizuoka	Baltimore	Lille	Southampton	Guangzhou
Hiroshima	Durham	<u>Londres</u>	<u>New-York</u>	Tianjin
Maebashi	Princeton	Strasbourg	Sheffield	Shenyang
Okayama	New-Haven	Rennes	Glasgow	Dalian
Tokai	Houston	Nancy	<u>Boston</u>	Xian
Kanazawa	San-Diego	Nice	Edimbourg	Hefei
Utsunomiya	Brooklyn	<u>New-York</u>	Leicester	Chengdu
<u>Beijing</u>	Atlanta	Nantes	Leeds	Kunming
Niigata	<u>Londres</u>	<u>Boston</u>	Nottingham	Jinan

*Comptages fractionnés (WNC), moyennes mobiles d'ordre 3

En **gras**, les agglomérations étrangères

Source: SCIEXP (articles, revues et lettres)

Tableau 30 – Les 15 premiers partenaires scientifiques des « villes globales » en nombre de collaborations en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Jusqu'ici nos résultats n'ont cessé de confirmer le caractère structurant du niveau national pour l'activité scientifique ce qui pourrait nous conduire à renoncer à l'idée de raisonner à un niveau aussi précis que le niveau urbain et à une échelle aussi petite que l'échelle mondiale pour revenir à une réflexion binaire de type nationalisation/internationalisation comme le suggérait Melin en 1999. Cependant, les figures ci-dessus indiquent, d'une part, que les structures interurbaines sont différentes d'un pays à

un autre et, d'autre part, qu'il se dégage des structures cohésives y compris à l'intérieur des pays. En effet, les groupes de couleur regroupés au sein de patatoïdes sur les représentations graphiques sont des groupes cohésifs mis en évidence grâce à la méthode des *Islands* (voir Chapitre 8.2.2). Cette méthode permet de détecter les groupes de villes qui entretiennent entre elles des relations dont l'intensité locale est plus forte que l'intensité de la plus forte des relations qu'elles entretiennent par ailleurs. D'une année à l'autre, la configuration de ces groupes se modifie légèrement. Dans la majorité des cas, les groupes obtenus correspondent à des pays entiers, mais il existe des groupes infranationaux aux États-Unis, en Allemagne et au Royaume-Uni (par exemple, l'Écosse ou la Californie constituent des groupes à part). Ces résultats s'expliquent très bien par le fait que ces trois derniers pays ont des organisations politiques plus fédérales.

Si les collaborations interurbaines ont volontiers progressé à l'intérieur des cadres nationaux entre 2000 et 2007, le réseau mondial s'est globalement densifié au cours de cette période, sans doute à l'avantage de certaines affinités intra et internationales. Ainsi, en se penchant uniquement sur les liens internationaux visibles sur les figures 64 et 65, on observe que l'intégration de la Chine et de la Corée du Sud au réseau mondial se fait par l'intermédiaire des États-Unis tandis que Paris fait figure d'intermédiaire entre le monde anglo-américain et le reste de l'Europe. En effet, même si les liens les plus intenses restent internes aux États, des relations tendent à s'instaurer entre toutes les villes du réseau mondial avec plus ou moins de force. D'une manière générale, il serait intéressant de savoir quels sont les espaces privilégiés de collaboration qui participent à déterminer l'évolution du réseau mondial. En effet, après avoir examiné la densification du réseau et son impact modéré sur la structure hiérarchique d'ensemble, il est temps de se pencher sur la structure géographique du réseau mondial telle qu'elle est sensiblement perceptible à travers les affinités entre villes. L'objectif de cette dernière section est de mettre en évidence le rôle conjugué des différents niveaux d'organisation sur la structure du réseau mondial. Cette question est d'autant plus essentielle qu'il existe des volontés à la fois scientifiques et politiques de renforcer la cohésion de certains sous-espaces. En particulier, le développement de la cohésion scientifique de l'espace européen est présenté comme un enjeu par les pouvoirs publics pour faire contrepoids à la force scientifique du continent nord-américain et à la puissance de plus en plus incontournable du continent asiatique. Au-delà même des stratégies politiques, il existe des affinités fortes entre les pays, qui ne sont pas seulement liées à leur position géographique mais aussi à la nature des liens qu'ils ont tissés par le passé. Dans certains cas, et en particulier dans le cas des anciens empires coloniaux, les « affinités » héritées du passé traversent aisément les océans.

2.2. Des aires géographiques de collaborations privilégiées

Depuis l'article fondateur de Frame et Carpenter en 1979, plusieurs auteurs ont montré qu'il existe des affinités scientifiques entre pays permettant de dégager de grandes aires de collaborations internationales à l'échelle mondiale. Cette famille de travaux fait ressortir des « clusters » ou groupes cohésifs de pays, c'est-à-dire des ensembles de pays qui collaborent intensément (Frame & Carpenter, 1979 ; Schubert & Braun, 1990 ; Okubo *et al.*, 1992 ; Glänzel & Schubert, 2005). Comme on dispose de données relationnelles à un niveau plus fin que pour ces précédents travaux, il est intéressant de voir dans quelle mesure on retrouve ce type de préférences à partir de l'analyse des relations interurbaines entre les 500 premières agglomérations ayant publié le plus dans le SCI Expanded en 2007. Puisque les résultats obtenus démontrent que les 500 agglomérations ont d'abord des relations privilégiées avec les agglomérations de leur pays d'appartenance, l'analyse est ensuite prolongée, dans un second temps, à l'intégralité des collaborations interurbaines entre pays, d'abord au niveau des pays européens puis au niveau de l'ensemble des pays du monde.

Pour détecter des groupes macro-régionaux au sein du réseau de collaboration des 500 premières agglomérations, on applique la « méthode de Louvain » (Blondel *et al.*, 2008). Elle consiste à optimiser une fonction de qualité, la modularité. La valeur de cet indicateur est d'autant plus grande que les collectifs de villes détectés bénéficient d'une forte cohésion interne et d'une propension plus faible à collaborer avec les villes des autres groupes. L'application de cette méthode ainsi que d'autres méthodes du même type comme la méthode VoS (Waltman *et al.*, 2010) au réseau des collaborations scientifiques entre les 500 premières agglomérations permet facilement de faire ressortir trois grandes aires : une aire de collaboration asiatique ; une aire de collaboration caractérisée par l'influence des villes américaines qui s'étend à quelques pays du Moyen-Orient ; et une grande aire de collaboration à laquelle participent les villes européennes mais qui comprend des villes d'autres continents comme l'Afrique, l'Océanie et l'Amérique Latine. Pour avoir un aperçu d'ensemble de ce résultat, une carte interactive de synthèse est accessible sur le site internet <http://www.Coscimo.net/> co-réalisé avec Laurent Jégou (voir Figure 47, page 249). Aussi, puisqu'en raison de la force des liens intra-nationaux, aucune ville ne se désolidarise des autres villes de son pays, le résultat obtenu à plusieurs étapes du partitionnement progressif du réseau des 500 premières agglomérations est résumé dans la Figure 66 qui liste la composition des groupes par pays. Cette figure se lit de haut en bas et indique la composition des groupes par pays, ce qui la rend plus lisible que s'il avait fallu lister l'ensemble des villes comprises dans chaque groupe.

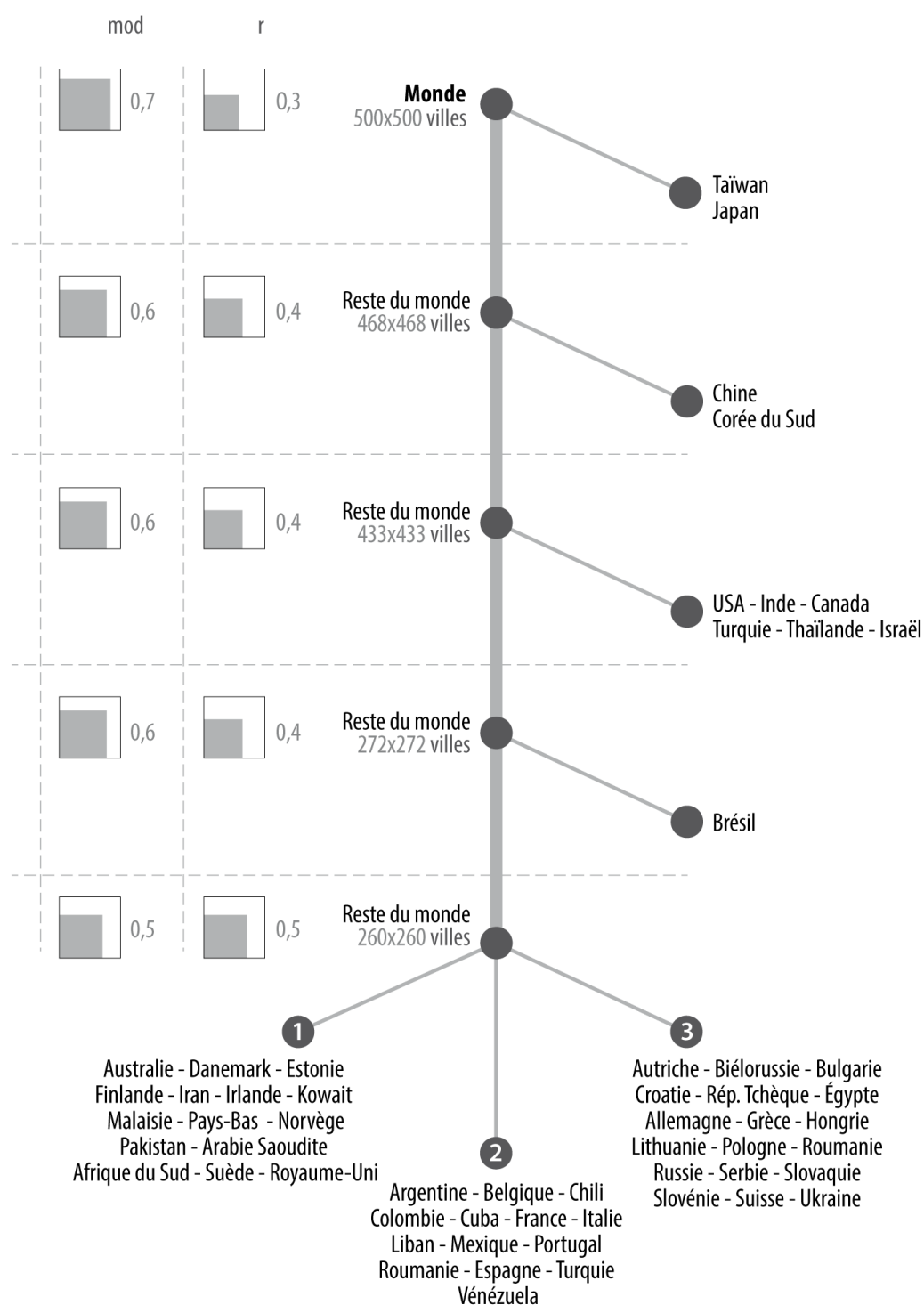
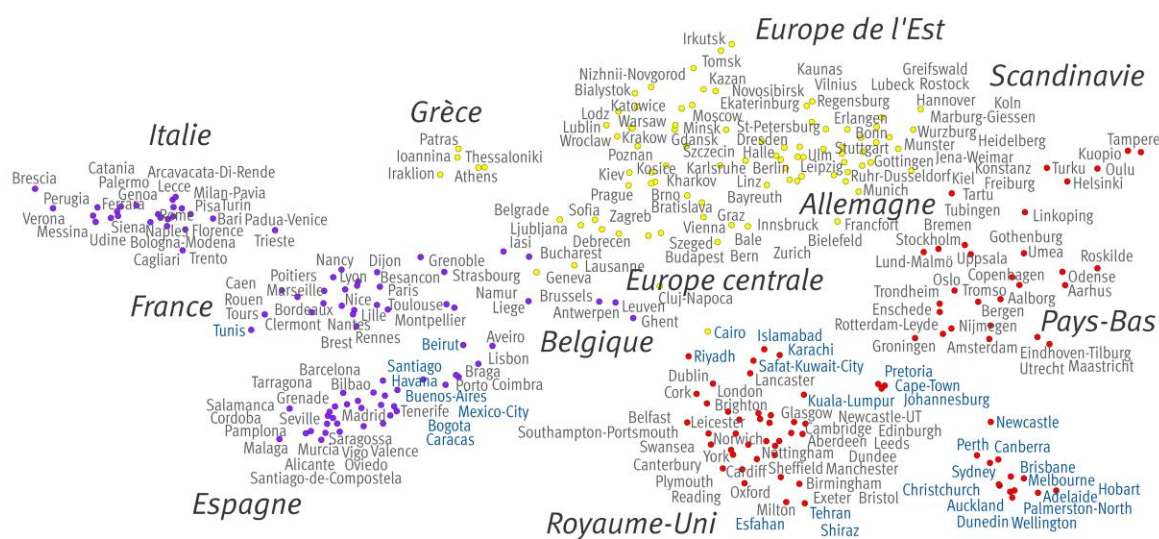


Figure 66 – Dendrogramme issu d'une détection de communautés de type Louvain appliquée pas à pas au réseau interurbain des collaborations scientifiques mondiales entre les 500 premières agglomérations en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Afin de proposer un résultat aussi représentatif que possible de la structure d'ensemble du réseau, la composante principale du réseau a été décomposée pas à pas en faisant augmenter progressivement la valeur du paramètre de résolution r fixé *a priori* (Figure 66). Plus la résolution est élevée, plus l'algorithme parvient à détecter de « communautés » et la taille des « communautés » diminue. Le réglage de ce paramètre équivaut en quelque sorte au réglage d'une focale. On démarre avec le plus petit r possible, c'est-à-dire en fixant une résolution tout juste assez haute pour que l'algorithme identifie au moins une coupure (en l'occurrence 0,3). À ce niveau de résolution, le groupe formé par les villes de Taiwan et du Japon se détache de l'ensemble des villes du monde. Cela signifie que l'intensité des échanges entre les villes de ces deux pays réunies est plus forte que l'intensité de leurs relations avec le reste du monde.

On augmente ensuite la valeur du paramètre de résolution pour détecter d'autres sous-groupes cohésifs au sein de la « communauté » formée par les agglomérations restantes (la composante principale appelée « reste du monde »). Pour commencer, le groupe des villes chinoises et coréennes apparaît suffisamment cohésif et indépendant du reste du monde pour en être détaché, puis c'est au tour du groupe des villes américaines accompagné de nombreuses villes étrangères sous influence américaine (161 agglomérations). Une fois le groupe américain exclu de la composante principale, l'ensemble des villes brésiliennes se révèle assez cohésif pour s'autonomiser du reste du monde. À ce stade, il reste une « communauté » assez large de 260 agglomérations qui comprend l'ensemble des 202 agglomérations européennes accompagné d'une cinquantaine de villes sous influence européenne. Cela suppose qu'à l'inverse des villes asiatiques, américaines et brésiliennes, les villes européennes ne constituent pas un ensemble dont les frontières soient faciles à établir à l'échelle mondiale. Avec un paramètre de résolution de 0,5 appliqué à ce réseau résiduel, on parvient toutefois à mettre en évidence trois sous-groupes de villes dont la composition est détaillée au pied de l'arbre (Figure 66). Mais, comme en témoigne la représentation graphique montrant la position des éléments de cet ensemble d'après leur degré mutuel de similarité (Figure 67), certaines villes de ce réseau se trouvent dans des positions telles qu'il serait tout aussi judicieux de les intégrer au groupe voisin.

Positionnement des villes d'après l'intensité de leurs collaborations scientifiques en 2007



Conception et réalisation: MM

Algorithme de visualisation: VoS

Source: SCI Expanded, (articles, review, letters)

Détection de communauté à l'aide de la méthode de Louvain (résolution 0,5, modularité 0,53), collaborations scientifiques 2006-2008

Figure 67 — Positionnement multiscalaire des villes européennes et des villes sous influence européenne d'après l'intensité de leurs collaborations scientifiques. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

D'après ce partitionnement, les villes françaises, italiennes, espagnoles et portugaises appartiennent au même groupe (en bleu) auquel il faut adjoindre les villes de 6 pays d'Amérique Latine (Argentine, Chili, Colombie, Cuba, Mexique, Vénézuëla) ainsi que Beyrouth et Tunis (le nom des villes appartenant à d'autres continents est en bleu sur la représentation). Les villes belges sont également dans ce groupe, ainsi que les villes roumaines (Bucarest, Iași), mais la représentation multiscalaire (obtenue à l'aide de l'algorithme de visualisation VoS) montre que ces dernières se trouvent à la frontière avec le groupe des villes d'Europe Centrale et de l'Est (en jaune). En fait, cette représentation permet de situer les villes les unes par rapport aux autres en fonction de leur affinité scientifique mesurée à l'aide d'un indice de similarité (Waltman *et al.*, *ibid.*). Elle montre la difficulté qu'il peut y avoir à trancher sur l'appartenance de certaines villes compte tenu de la faible distance relationnelle qui les sépare. Bien qu'elle soit intéressante, cette division de l'Europe en trois sous-groupes ne devrait pas être sur-interprétée. La qualité du résultat n'est que légèrement supérieure à son seuil de significativité (la modularité est égale à 0,5). En somme, alors que toutes les villes représentées forment un groupe cohésif à l'échelle mondiale, les trois sous-groupes détectés à l'intérieur de ce large ensemble sont loin d'être parfaitement indépendants.

Il ressort de cette représentation que l'affinité réciproque des 202 agglomérations de l'Union Européenne à 28 n'est pas suffisamment forte pour qu'elles se détachent mécaniquement de l'ensemble du réseau des agglomérations les plus publiantes du monde en 2007. En effet, la représentation multiscalaire indique que les villes du Royaume-Uni se trouvent bien plus éloignées des villes de Grèce et d'Italie en terme d'affinité scientifique qu'elles ne le sont des villes d'Afrique du Sud, d'Australie et des quelques métropoles du Moyen-Orient, qui partagent un langage et un passé colonial avec le Royaume-Uni. Il semblerait donc, à la vue de cette représentation, que l'Espace Européen de la Recherche (EER) ne se comporte pas comme un niveau d'organisation hyper-structurant dans la mesure où il n'éclipse pas le poids des proximités d'ordre linguistique ou historique qui rapprochent les agglomérations de l'Union de certaines agglomérations extérieures à l'Union. C'est encore plus net lorsqu'on tient compte de l'ensemble des relations interurbaines du monde. Dans ce cas, on voit très clairement apparaître des relations privilégiées qui transcendent les frontières continentales comme entre la France et l'ensemble des pays africains francophones (Figure 68). Pour autant, le renforcement de l'EER est souvent présenté comme un enjeu lié à la globalisation.

En effet, plus d'une trentaine de pays européens partagent désormais la volonté de faire exister l'EER. Cette ambition, déjà présente au sortir de la Seconde Guerre mondiale (Euratom est encore là en 2015 pour en témoigner), et au moment du lancement des PCRD (Programmes Cadre de Recherche et de Développement) dans les années 1980, a été réaffirmée dans le cadre de la Stratégie de Lisbonne au début des années 2000. Pour renforcer la cohésion interne de l'espace communautaire, les chercheurs européens ont donc à leur disposition de nombreux dispositifs destinés à accroître leurs collaborations. Puisque ces initiatives facilitent les communications, il y a même des pays européens extérieurs à l'UE qui s'y intéressent. C'est pourquoi, d'après le site internet de l'*European Research Area and Innovation Committee* (ERAC)⁹⁸, l'EER se compose aujourd'hui des 28 pays membres de l'UE et de 13 pays « observateurs ». La liste des « observateurs » se décompose en « pays associés » (l'Albanie, la Bosnie Herzégovine, les Iles Féroé, l'Ancienne République Yougoslave de Macédoine, la Moldavie, le Monténégro, la Serbie, la Suisse et la Turquie), « pays de l'Association Européenne de Libre Échange » ou EFTA (L'Islande, le Liechtenstein et la Norvège) et en « autres pays » (Israël). Mais, lorsqu'on considère l'ensemble des collaborations scientifiques interurbaines et donc qu'on intègre la participation des agglomérations extérieures aux 500 agglomérations les plus publiantes en 2007, on constate que le comportement de la totalité des agglomérations européennes en matière de collaborations scientifiques s'est très peu modifié entre 2000 et 2007 (Tableau 31).

⁹⁸ Source: URL: <http://erawatch.jrc.ec.europa.eu/>

Type de collaborations	% en 2000*	% en 2007*	Nombre de collaborations scientifiques en 2000*	Nombre de collaborations scientifiques en 2007*
A l'intérieur de l'UE à 28	57,1	57,3	65539,4	98280,3
UE à 28 avec le reste du monde	37,6	37,0	43116,0	63512,3
UE à 28 avec les pays partenaires de l'UE**	5,3	5,6	6119,0	9659,5
Total des collaborations avec au moins une ville de l'UE à 28	100,0	100,0	114774,5	171452,1
Total des collaborations scientifiques mondiales	49,6	45,3	231589,4	378098,2462

*Comptage fractionné (WNC), moyenne mobile sur 3 ans.

**Pays partenaires de l'UE28 dans le cadre de l'EER: Albanie, Bosnie Herzégovine, Îles Féroé, Macédoine, Moldavie, Monténégro, Serbie, Suisse, Turquie, l'EFTA (Islande, Liechtenstein et la Norvège), et Israël

Nombre total de publications scientifiques (articles, recensions et lettres) de l'UE28 en 2007*: 359 464,4

Source: SCIEp (articles, recensions et lettres)

Tableau 31 – L'évolution de la répartition des collaborations scientifiques de l'Union Européenne à 28. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Tandis que le nombre total des collaborations de l'UE a progressé de 50 % entre 2000 et 2007 (contre une croissance de plus de 63 % pour l'ensemble des collaborations mondiales), le profil des collaborations est resté pratiquement inchangé. Tout au long de la période retenue, 57 % de l'ensemble des collaborations de l'UE sont intra-communautaires (intra-UE) tandis que 43 % ont lieu avec le reste du monde, parmi lesquelles un peu plus de 5 % avec les chercheurs des pays « observateurs » ou partenaires de l'UE (intra-EER). On note tout au plus une légère progression en faveur de la cohésion de l'EER entre les deux dates, puisque la part des liens européens vers les pays extérieurs à l'EER (ni membres, ni observateurs) a diminué de 0,6 point. Comme cette évolution est minime, il est plus prudent de conclure à la stabilité. Mais cette stabilité mérite d'être interprétée à la lumière du contexte global précédemment étudié.

Alors que la déconcentration des activités scientifiques s'est traduite par l'intégration au système mondial de nouveaux espaces de production, notamment asiatiques (les collaborations européennes représentaient près de la moitié des collaborations mondiales en 2000 alors qu'elles n'en représentent plus que 45 % en 2007), la stabilité du comportement collaboratif de l'UE semble traduire une indifférence à cette nouvelle donne. Elle suggère que les chercheurs européens n'ont pas jeté leur dévolu sur les nouvelles opportunités de collaborations qui se sont présentées mondialement. Bien que le reste du monde se soit « agrandi », ces derniers n'ont pas cherché à s'y connecter davantage que par le passé. Si cela peut venir des incitations qui ont pesé pour coopérer à l'intérieur de l'EER, il convient aussi d'envisager le rôle de l'« agrandissement » de l'espace européen lui-même. Pour vérifier dans quelle mesure cette attitude est partagée

par d'autres grands espaces de collaborations, nous allons procéder à une décomposition du réseau de l'ensemble des collaborations interurbaines du monde en 2000. En se fondant sur le résultat obtenu, nous allons mesurer l'évolution de la répartition des collaborations au sein et entre les grands espaces de collaborations à l'échelle mondiale.

Pour délimiter le périmètre des grands espaces de collaboration, on s'appuie sur la répartition des collaborations interurbaines entre l'ensemble des pays n'appartenant pas à l'EER. Le réseau des collaborations scientifiques ainsi obtenu peut se décomposer en une dizaine de petites aires de collaborations. En confrontant les résultats obtenus à l'aide de la méthode de Louvain et de la méthode VoS (la seconde normalise davantage par la taille que la première et détecte un plus grand nombre de communautés) et en faisant varier le paramètre de résolution, sept grandes régions sont définies :

- Le monde russe regroupant la Russie ainsi que des pays frontaliers pour la plupart anciennement soviétiques comme les pays du Caucase. Ce groupe intègre aussi l'Iran qui est difficile à classer car elle entretient, parallèlement, un nombre relativement important de relations scientifiques avec les pays du Commonwealth (le Canada en particulier) et les États-Unis.

- L'Amérique du Nord, Centrale et Caraïbe qui comprend les États-Unis, le Canada, plusieurs micro-États des Caraïbes et d'Océanie ainsi que le sous-ensemble des pays de la cordillère centrale du continent américain. Ce dernier sous-ensemble de pays, dont la cohésion scientifique interne est claire, montre autant de signes d'affinité scientifique avec l'Amérique du Nord qu'avec les pays d'Amérique Latine (incluant le Mexique).

- Le groupe océanique constitué de l'Australie, la Nouvelle-Zélande, l'Asie du Sud-Est, les îles de Mélanésie et de Polynésie ainsi que le Sri Lanka. Plusieurs pays d'Asie du Sud-Est qui sont également proches du monde asiatique en relèvent, mais le lien fort qu'ils partagent avec l'Australie justifie leur inclusion dans cet ensemble.

- Le monde asiatique regroupe l'Inde, le Bangladesh, le Népal, la Chine, le Japon, la Corée du Sud et du Nord (cette dernière n'entretient qu'un très faible nombre de relations et n'a que la Chine pour partenaire) ainsi que Taiwan. Taiwan et la Chine cultivent, par ailleurs, un fort attachement pour Singapour qui se trouve dans le groupe océanique.

- Le monde arabe intègre les pays du Maghreb et du Moyen Orient.

- L'Amérique Latine se compose des pays du Sud du continent américain ainsi que du Mexique.

- L'Afrique sub-saharienne regroupe un grand nombre de pays du Centre et du Sud du continent africain.

La composition exacte des groupes est mise en évidence sur la Figure 68. Cette figure montre la position de l'ensemble des pays du monde d'après leur degré d'affinité scientifique (mesuré à partir de la matrice des collaborations interurbaines). La position

des pays de l'EER sur cette représentation montre que s'ils avaient été pris en compte pour le partitionnement du monde, la teneur des groupes aurait été un peu différente. En particulier, l'Espagne serait comprise dans le groupe des pays hispano-américains et la France dans celui des pays africains francophones. Comme nous l'avons vu précédemment, il aurait été difficile de faire émerger l'EER à partir de ces seules données de collaboration scientifique. Ceci étant, il est intéressant de mettre en regard l'évolution de la cohésion interne à l'EER avec l'évolution de la cohésion interne aux autres grands espaces de collaboration.

Positionnement multiscalaire des pays dans le réseau des collaborations scientifiques internationales en 2000

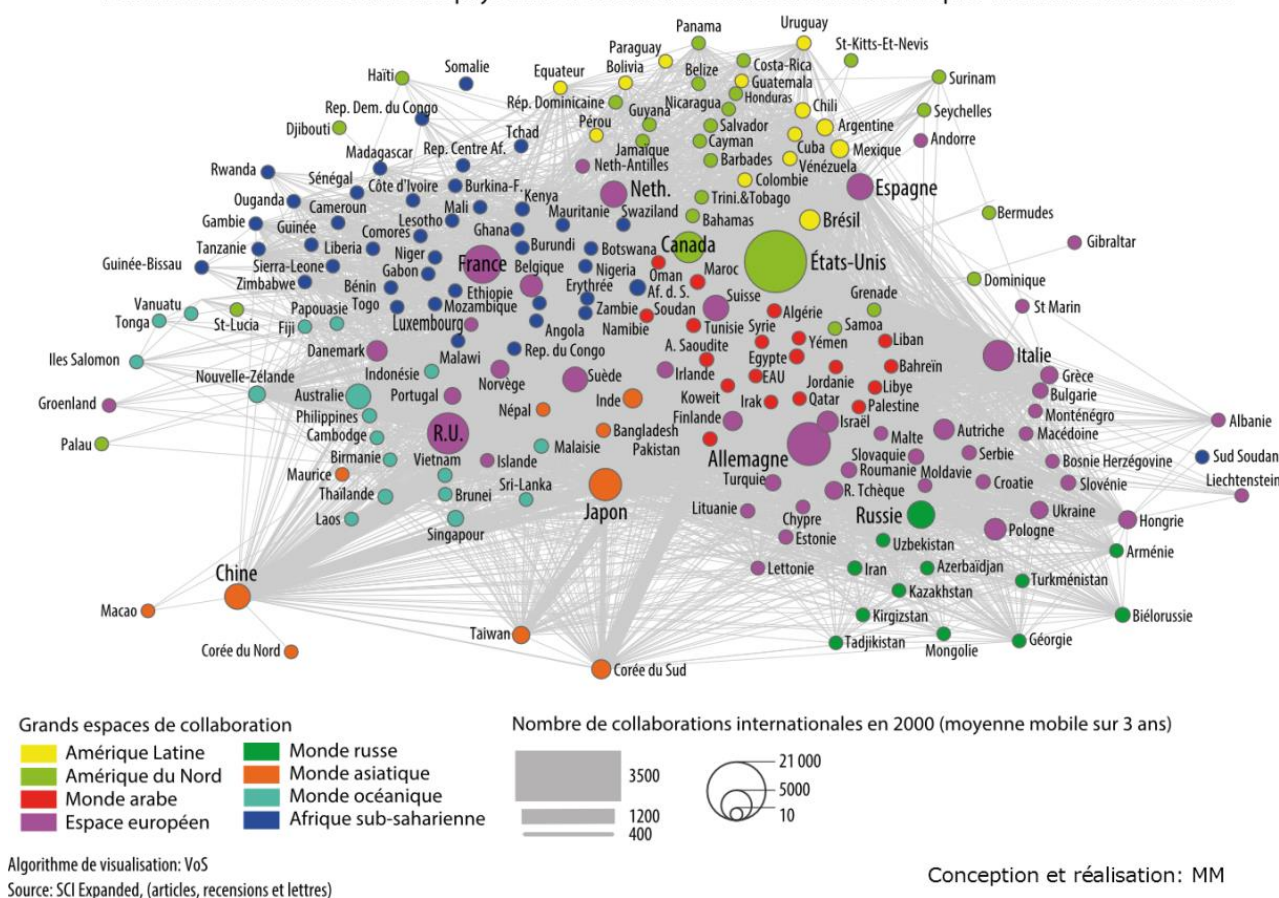


Figure 68 – Positionnement multiscalaire des pays dans le réseau des collaborations scientifiques mondiales en 2000. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Dans pratiquement tous les groupes de collaboration sauf dans celui d'Amérique du Nord, la part des relations extérieures au groupe a diminué entre 2000 et 2007. Cela s'explique par l'augmentation plus rapide de la part des collaborations infranationales et infrarégionales. Ce résultat suggère que les groupes de pays élaborés à partir de la réparti-

tion des collaborations scientifiques en 2000 sont toujours structurants en 2007. Il n'y a que dans le monde russe et en Amérique Latine que la cohésion semble s'être atténuée au profit des collaborations intra-nationales. En fait, les collaborations régionales ont progressé dans ces deux ensembles (+ 74 % en Amérique Latine et + 31 % dans le monde russe), mais beaucoup moins vite que les relations intra-nationales (+ 174 % en Amérique Latine et + 112 % dans le monde russe). Dans le monde arabe, l'Afrique sub-saharienne, le monde asiatique et océanique, les taux d'augmentation sont impressionnants à tous les niveaux. Le fait qu'ils le soient tout particulièrement à l'intérieur et entre les pays de chaque groupe explique la place plus grande qu'occupe ce type de collaboration au détriment des relations entre grands espaces.

Espace Européen	2000*	2007*		Monde russe	2000*	2007*	
Liens intra national (%)	34,5	35,6	↗	Liens intra national (%)	17,2	26,5	↗
Liens intra Europe (%)	26,8	26,4		Liens intra monde russe (%)	2,0	1,9	
Liens avec le RDM** (%)	38,7	38,0		Liens avec le RDM** (%)	80,8	71,6	↘
	100	100			100	100	
Nombre de publications	122994	186408		Nombre de publications	8917	12207	
Amérique du Nord	2000*	2007*		Monde océanique	2000*	2007*	
Liens intra national (%)	51,9	50,6	↘	Liens intra national (%)	20,1	19,7	
Liens intra Am. Du N. (%)	4,1	4,1		Liens intra Océanie (%)	4,7	5,1	↗
Liens avec le RDM** (%)	44,0	45,3	↗	Liens avec le RDM** (%)	75,2	75,2	
	100	100			100	100	
Nombre de publications	98043	147374		Nombre de publications	10201	20182	
Monde asiatique	2000*	2007*		Monde arabe	2000*	2007*	
Liens intra national (%)	55,2	57,0	↗	Liens intra national (%)	11,2	14,5	↗
Liens intra Asie (%)	4,5	5,2	↗	Liens intra monde arabe (%)	5,8	7,0	↗
Liens avec le RDM** (%)	40,3	37,8	↘	Liens avec le RDM** (%)	83,1	78,5	↘
	100	100			100	100	
Nombre de publications	43634	89710		Nombre de publications	2687	5705	
Amérique Latine	2000*	2007*		Afrique sub-saharienne	2000*	2007*	
Liens intra national (%)	27,5	38,5	↗	Liens intra national (%)	14,9	16,2	↗
Liens intra Am. Lat. (%)	5,9	5,2		Liens intra Afrique sub-s. (%)	3,7	5,4	↗
Liens avec le RDM** (%)	66,6	56,3	↘	Liens avec le RDM** (%)	81,4	78,4	↘
	100	100			100	100	
Nombre de publications	9458	18530		Nombre de publications	2652	5089	

*Comptage entier fractionné (WNC), moyenne mobile sur 3 ans. Source: SCI Expanded (articles, recensions et lettres)

**RDM = Reste du Monde

Tableau 32 — L'évolution de la répartition des collaborations scientifiques entre grands espaces de collaboration (2000-2007). Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

Cependant, même si elles ont évolué moins vite que les relations internes aux grands espaces de collaboration, les relations entre grands espaces ont également augmenté en valeur absolue. Leur augmentation et les ordres de grandeur varient beaucoup selon les espaces, allant de + 3 % pour le nombre de collaborations entre le monde russe et l'Amérique Latine (de 144 à 149 co-publications) à + 383 % pour le nombre de collaborations entre le monde russe et le monde arabe (de 13 à 63 co-publications). La Figure 69

permet de rendre compte de l'intensité de ces évolutions (selon l'intensité de rouge) ainsi que des ordres de grandeur auxquels ces augmentations font référence en 2007 (l'épaisseur des flux pour les relations entre espace, et la taille des cercles pour les liens réflexifs).

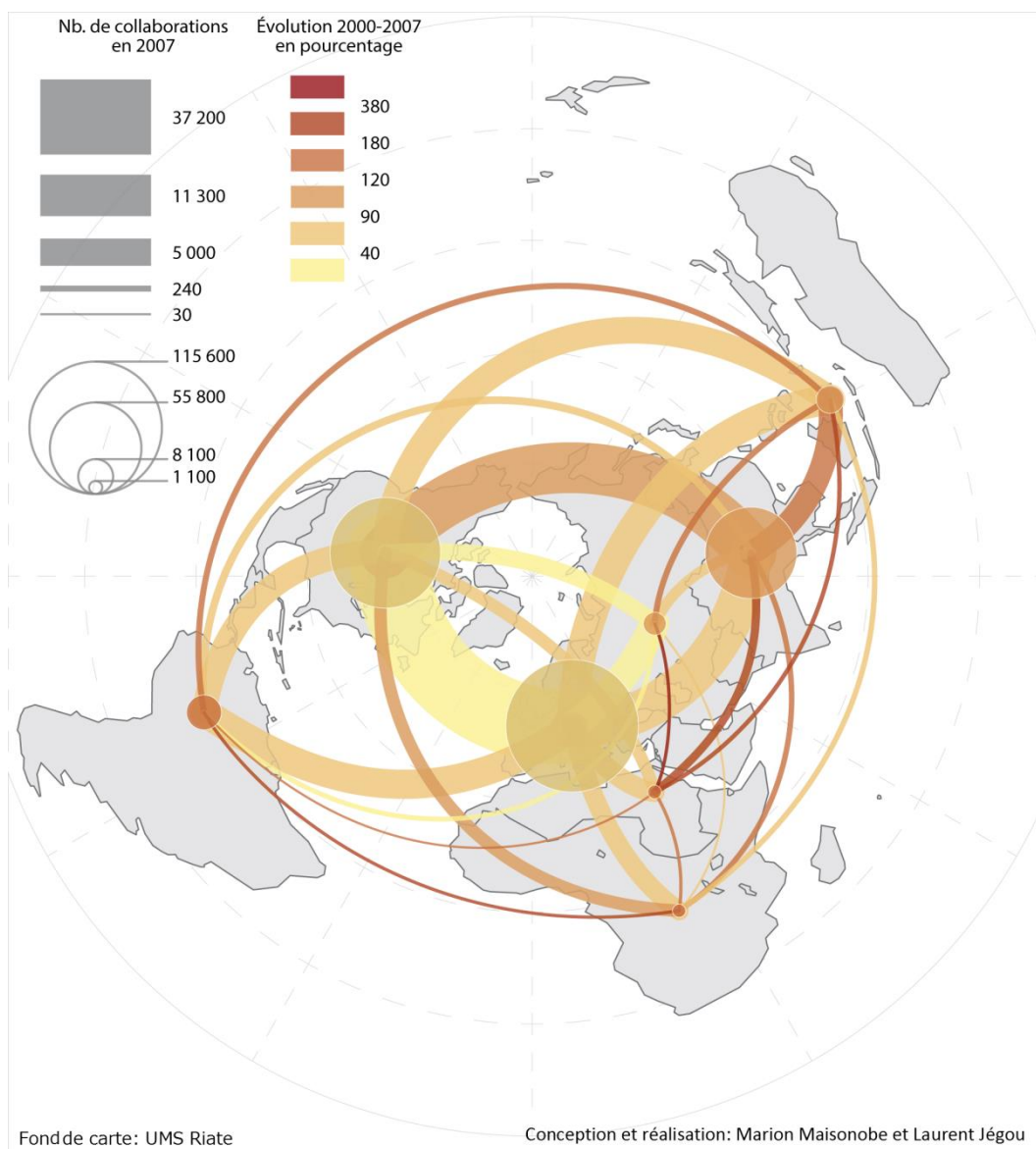


Figure 69 – Carte de l'évolution des collaborations scientifiques entre grands espaces de collaborations mondiaux (2000-2007). Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans

On observe que les évolutions les plus fortes concernent les espaces autrefois les plus périphériques au détriment des relations avec l'Europe et l'Amérique du Nord. Le monde asiatique apparaît être l'espace ayant le plus participé à la croissance des relations puisqu'il cumule d'importants volumes de collaborations et d'importants taux de crois-

sance. Même si les volumes sont moindres, les liens entre l'Amérique Latine, le monde océanique, le monde russe, le monde arabe et l'Afrique sub-saharienne augmentent plus vite que leurs relations respectives avec l'Europe et l'Amérique du Nord. Cela signifie que de plus en plus de relations se nouent entre les espaces « non-hégémoniques » du système scientifique mondial.

Ce résultat va à l'encontre de l'idée véhiculée par Caroline Wagner dans un article co-écrit avec Loet Leydesdorff (Leydesdorff & Wagner, 2008) selon laquelle, malgré la croissance globale des activités de collaboration scientifique, la cohésion du réseau scientifique mondial augmenterait au sein d'un « cœur » d'une quinzaine de pays centraux pouvant s'appuyer sur de forts systèmes nationaux au détriment de l'intégration d'un nombre important de pays qui seraient de plus en plus périphériques. Il conforte, en revanche, le point de vue défendu par Ponomariov et Toivanen à partir du cas brésilien : « le potentiel de recherche à la fois domestique et Sud-Sud joue un rôle croissant dans les pays développés et les systèmes émergents d'innovation économique en même temps que ces derniers se renforcent. »⁹⁹ (Ponomariov & Toivanen, 2014).

Ainsi, ce dernier résultat montre encore qu'à l'encontre des idées reçues sur la globalisation, la géographie des activités scientifiques de cette dernière décennie n'a pas été caractérisée par un renforcement global des inégalités spatiales dicté par une « main invisible » obéissant aux lois d'un « marché scientifique » privilégiant une poignée de villes ouvertes sur l'international et appelées à structurer le système mondial en s'appuyant sur leur fort système de recherche national. Au contraire, les données que nous venons de décrire montrent que tous pays confondus, il n'y a pas eu de franche tendance à la nationalisation ou à l'internationalisation des collaborations au cours de la dernière décennie, mais plutôt une stabilité des comportements collaboratifs. En dépit de cette stabilité, le réseau mondial s'est densifié entre les deux périodes étudiées et de moins en moins de membres en sont isolés. Les liens progressant le plus vite ont été les liens intra-nationaux tandis que les villes asiatiques ont détrôné plusieurs villes d'Europe de l'Est et du Nord en haut de la hiérarchie. Enfin, les structures relationnelles perceptibles à travers les affinités entre villes et entre pays d'après l'intensité de leurs relations ont permis de faire émerger des familles de villes et de pays. Ces résultats suggèrent que la croissance des collaborations mondiales, qualifiée parfois de « globalisation », ne se fait ni au détriment du renforcement des systèmes nationaux, ni au détriment de celui des grandes aires géographiques, linguistiques, culturelles ou civilisationnelles mondiales.

Ce qu'il faut retenir de l'ensemble de ces résultats exploratoires, c'est donc que la densification des collaborations scientifiques se déroule dans un contexte mondial dont la

⁹⁹ L'original : « domestic and South-South research capacities play an increasing role in developing country and emerging economy innovation systems as they gain momentum » (Ponomariov & Toivanen, 2014).

structuration est le fruit de lointains héritages, parmi lesquels le caractère structurant des cadres nationaux n'est pas en reste. Les chercheurs, contrairement à une imagerie commune ne se comportent pas comme des électrons libres dans un espace scientifique insensible au poids des frontières géographiques, mais travaillent au sein d'espaces organisés, et obéissent à des logiques sociales d'appariement multiples, au rang desquelles « l'attachement préférentiel » (l'attraction pour les plus « gros » ou « riches » en nombre de relations) n'est qu'une option parmi d'autres.

Cette multiplicité des logiques d'appariement et des niveaux d'organisation explique qu'il soit si difficile de délimiter des espaces significativement cohésifs à partir de la répartition des liens de collaboration. Sauf dans le cas des pays d'organisation fédérale, les espaces qui se révèlent être les plus cohésifs sont les espaces nationaux. Comme il est quelque peu illusoire de chercher à positionner chaque entité spatiale dans un seul espace de collaboration à la fois, de récentes méthodes ont été développées qui autorisent les appartenances multiples. C'est le cas de la méthode OSLOM (Lancichinetti, Radicchi, Ramasco, & Fortunato, 2011), qui permet de mettre simultanément en évidence de très grands groupes et des groupes cohésifs d'échelle nationale voire régionale. Il s'agit d'une méthode de détection de « communautés chevauchantes ». Appliquée au réseau des 500 agglomérations, elle permet de faire ressortir à la fois des groupes nationaux très cohésifs comme le groupe des agglomérations néerlandaises ; quelques groupes infranationaux très cohésifs comme le groupe des agglomérations écossaises ; de petites familles de pays comme les agglomérations scandinaves accompagnées de Kaunas et Vilnius (Lituanie) ; et de plus grands groupes comme celui associant les villes allemandes avec les villes ukrainiennes et russes. Enfin, l'ensemble des villes françaises et Londres se retrouvent dans plusieurs groupes à la fois : un groupe qui les connecte aux villes italiennes ; un groupe qui les connecte au reste des villes francophones (belges, suisses, québécoises, et tunisiennes) ; et un groupe qui les connecte au reste des villes hispaniques et lusophones.

Cette approche mériterait sans doute d'être explorée davantage, mais il reste à estimer comment exploiter la richesse des résultats ainsi obtenus. En effet, les résultats des classifications gagnent vraiment en intérêt à la lumière des interprétations. Or, à cet égard, les facteurs « macro-géographiques » et « macro-historiques » ne suffisent pas car plus on se rapproche des objets scientifiques, plus on constate qu'il existe des raisons internes à la science. En particulier, d'une discipline à l'autre, la géographie des activités de production scientifique diverge à la fois par ces évolutions et par sa structuration (en témoignent les deux tableaux suivants qui montrent la plus ou moins grande propension des recherches à s'internationaliser et se collectiviser selon les disciplines). Au-delà, plus on cherche à « contrôler » par l'effet pays, la langue, le poids des unités spatiales et la distance qui les sépare, plus on est en mesure de voir apparaître des relations « aberrantes », c'est-

à-dire impossible à expliquer à moins d'entreprendre des recherches plus poussées sur les observations. Elles peuvent bien sûr avoir des causes d'ordre géographique ou historique mais elles tiennent aussi souvent à des réalités très fines qui sont celles de la science en pratique.

D'après le Tableau 33 et le Tableau 34, les publications scientifiques à une seule adresse décroissent en proportion pour toutes les disciplines sauf la Physique, les Sciences de l'Univers et les Sciences de l'ingénieur. Pour toutes les autres disciplines, les évolutions suivent les tendances globales et s'inscrivent dans le prolongement des évolutions déjà mesurées par Gingras *et al.* en 2002 (*op. cit.*).

Part en 2007* (%) et évolution entre 2000* et 2007*						
Répartition des publications	Bio. Fonda.	Médecine	Chimie	Physique	Bio. Appliquée	Sc. de l'Univers
Une seule ville, une seule adresse	34,1 ↘	33,3 ↘	47,8 ↘	51,1 ↗	39,7 ↘	65,9 ↗
Une seule ville, plusieurs adresses	25,8 ↗	31 ↗	19,2 ↗	14,7 ↗	18,4 ↗	11 ↘
Plusieurs villes du même pays	19,2 ↗	20,1 ↗	16,7 ↗	14,1 ↗	21,5 ↗	11,3 ↘
Plusieurs villes, plusieurs pays	20,9 ↗	15,7 ↗	16,2 ↗	20,1 ↘	20,5 ↗	11,8 ↘
Total (%)	100	100	100	100	100	100
Nombre total de publications en 2007*	201183,3	345387,3	197166,7	172604	97582,7	24606,7

*Normalized counting (WNC), moyenne mobile sur 3 ans. Source: WoS (Nombre total d'articles, revues et lettres)

**Ecart entre 2000 et 2007: de 0 à 3 points d'écart (1 flèche); de 3 à 15 points d'écart (2 flèches); au-delà de 15 points d'écart (3 flèches)

Tableau 33 – Structure géographique de la production scientifique mondiale par discipline.

Source : Web of Science (articles, reviews, letters), moyenne mobile sur 3 ans

En fait, la Physique est dans un cas particulier d'augmentation des collaborations intra-nationales au détriment des collaborations internationales. Les Mathématiques sont aussi dans ce cas-là mais, inversement, la part des articles de Mathématiques à une seule adresse est passée de 60 à 55 % entre 2000 et 2007 (Tableau 34). Dès lors, constatant une tendance au repli domestique des Mathématiques et de la Physiques ainsi que des Sciences de l'Univers et de l'Ingénieur, on peut faire l'hypothèse d'une contrainte spécifique de l'instrumentation sur les collaborations, dans la mesure où les sciences les plus instrumentées sont celles qui se sont le moins ouvertes à l'international entre 2000 et 2007 ; mais il faudrait des études plus approfondies pour pouvoir le confirmer et le préciser.

Part en 2007* (%) et évolution entre 2000* et 2007*					
Répartition des publications	Sc. de l'ingé.	Mathématiques	Multi-disciplin.	Arts & Human.	Sc. Sociales
Une seule ville, une seule adresse	65,9 ↗	55,4 ↘	55,5 ↘	63,6 ↘	54,9 ↘
Une seule ville, plusieurs adresses	11 ↘	10 ↗	17 ↗	10,5 ↗	13,1 ↗
Plusieurs villes du même pays	11,3 ↘	13,2 ↗	13,7 ↗	15,5 ↗	19,5 ↗
Plusieurs villes, plusieurs pays	11,8 ↘	21,4 ↘	13,9 ↗	10,4 ↗	12,6 ↗
Total (%)	100	100	100	100	100
Nombre total de publications en 2007*	235140,3	47066,3	23771,3	62253,7	69362,7

*Normalized counting (WNC), moyenne mobile sur 3 ans. Source: WoS (Nombre total d'articles, revues et lettres)

**Ecart entre 2000 et 2007: de 0 à 3 points d'écart (1 flèche); de 3 à 15 points d'écart (2 flèches); au-delà de 15 points d'écart (3 flèches)

Tableau 34 – Structure géographique de la production scientifique mondiale par discipline.

Suite. Source : Web of Science (articles, reviews, letters), moyenne mobile sur 3 ans

L'existence de logiques internes à l'activité scientifique justifie que nous poursuivions des recherches sur des sous-ensembles de publications préalablement bien définis qui puissent se révéler représentatifs de « communautés » épistémiques ou de pratique (Meyer & Molyneux-Hodgson, 2011). Comme nous voulons l'illustrer à travers l'exemple de la spécialité de la réparation de l'ADN, les comportements de collaboration varient non seulement au niveau des disciplines mais aussi des spécialités scientifiques. Se rapprocher de la fabrication des contenus scientifiques permet de comprendre combien la géographie de la production scientifique est le résultat d'une conjonction de facteurs que l'on peut tenter d'appréhender en considérant le rôle et les contraintes des trajectoires individuelles, des trajectoires organisationnelles, mais aussi des problèmes et des pratiques de recherche. De même, les connexions qui se mettent en place entre les différents lieux de la pratique scientifique résultent de la mobilisation de logiques d'appariement et de dispositifs de coordination variés, à différents niveaux, dont le niveau individuel et le niveau organisationnel. Étudier précisément, par le biais d'une approche géographique, la mise en place de collaborations scientifiques à l'échelle mondiale dans l'optique de développer une nouvelle spécialité (la réparation de l'ADN de 1965 à nos jours), et, au sein même de cette spécialité, de résoudre un problème de recherche « à cheval » sur deux spécialités en biologie moléculaire (la question du lien entre transcription et réparation de l'ADN de 1990 à nos jours) est l'objet de la toute dernière partie de ce travail.

PARTIE 3.
UNE SPÉCIALITÉ SCIENTIFIQUE,
« LA RÉPARATION DE L'ADN »

Le traitement géographique des données du *Web of Science* a permis de mettre en évidence plusieurs caractéristiques de la science contemporaine. Plus spécifiquement, il a été l'occasion d'analyser une portion entière de la science produite dans le monde, celle qui répond aux standards « internationaux » tels qu'ils sont interprétés par les gestionnaires de la base de données. Par ce biais et en nous soumettant à un exercice de ventilation et d'agrégation des données à différents niveaux géographiques, nous avons trouvé la preuve qu'au cours de la dernière décennie, les publications scientifiques proviennent d'un nombre toujours plus important de lieux. Ce phénomène s'explique par les politiques nationales de décentralisation des activités de recherche et d'enseignement supérieur menées dans un nombre important de pays aux lendemains de la Seconde Guerre mondiale mais également par la participation volontaire et active d'États autrefois peu impliqués dans le système comme la Chine et les Émirats Arabes Unis. Comme le signalait Jean Gottmann en 1973, nous avons affaire à un système mondial qui s'agrandit et cette expansion requiert davantage de dispositifs de coordination.

En théorie, les échanges scientifiques sont d'autant plus nombreux que la division du travail scientifique est approfondie et donc, que la science est spécialisée. À mesure que le nombre de scientifiques augmente, le nombre de relations possibles augmente à cela près qu'apparaissent des frontières internes aux activités scientifiques. Pour Derek De Solla Price, la spécialisation scientifique est un effet de la croissance de la science (dont témoigne l'inflation de littérature scientifique), qui est elle-même un effet de l'augmentation du nombre de participants à l'entreprise scientifique (Solla Price, 1962 cité par Polanco, 1995). Le phénomène de spécialisation est structurant pour l'activité scientifique puisque la spécialité est un niveau d'organisation sociale des chercheurs en vue de la résolution de problèmes. Les travaux portant sur l'organisation de la science ont donc souvent insisté sur ce niveau d'organisation qui correspond à un idéal : le « collège invisible ». Parmi les cercles professionnels auxquels les chercheurs peuvent appartenir, les « collèges invisibles » sont *a priori* ceux dont l'étendue géographique est la plus diffuse. L'interconnaissance et la communication entre les scientifiques est supposée optimale à ce niveau. Alors que les scientifiques ne sont pas incités à connaître et communiquer avec tous les membres de leur discipline, ni avec tous les membres de leur laboratoire, l'institution attend d'eux qu'ils s'intègrent à une spécialité dont ils sont supposés connaître les membres et les travaux.

Pour mieux saisir les ressorts de l'organisation des activités scientifiques à l'échelle mondiale, il est intéressant donc de voir ce qui se passe au niveau d'une spécialité scientifique. L'étude de cas réalisée dans cette perspective porte sur la spécialité de la réparation de l'ADN. Plusieurs processus ont été interrogés : l'émergence, le développement et l'organisation de cette spécialité à plusieurs niveaux et à différents moments.

CHAPITRE 10. L'ÉMERGENCE ET LA DIFFUSION D'UNE SPÉCIALITÉ SCIENTIFIQUE : LA RÉPARATION DE L'ADN

Comment expliquer la géographie actuelle des recherches sur la réparation de l'ADN ? Comment spatialiser le phénomène d'émergence et de diffusion des recherches en ce domaine ? Dans quelle mesure l'expansion géographique de la spécialité est-elle associée aux trajectoires individuelles de ses pionniers ? Que nous apprend l'histoire de ce collectif sur la dynamique du système scientifique mondial ?

L'approche géographique vient ici se substituer à une approche plus classique de sociologie des sciences pour interroger l'émergence, l'organisation et l'expansion d'une spécialité à l'échelle mondiale. Cela présente un triple intérêt :

- saisir la diversité des logiques d'attachement qui justifient la mise en relation de scientifiques localisés en des lieux différents ;
- voir dans quelle mesure l'organisation et la répartition des chercheurs est dépendante de la structure du système scientifique mondial ;
- comprendre comment les chercheurs contribuent, en s'organisant au niveau d'une spécialité de recherche, à la structuration du système scientifique mondial.

Dans ce chapitre, nous commençons par revenir sur les études sociologiques de spécialités scientifiques. En s'appuyant sur ces travaux, nous préconisons l'adoption de méthodes mixtes, à la croisée du qualitatif et du quantitatif.

Pour commencer, nous abordons les enjeux méthodologiques associés à la constitution d'un corpus bibliographique consacré à l'étude d'une spécialité scientifique. L'analyse bibliométrique permet ensuite de comparer la géographie de la spécialité à deux stades de recherche : le stade d'émergence et un stade ultérieur, celui de la maturité. Et, pour aller plus loin dans la compréhension du phénomène d'émergence, une étude de la trajectoire des pionniers de la spécialité est réalisée. Elle offre la possibilité de saisir les logiques de diffusion du front de recherche et d'identifier les ressorts de son expansion géographique. Cette étude est enrichie d'archives, de témoignages écrits et oraux des membres de la communauté.

1. L'étude d'une spécialité scientifique

« The first attempts to write the history of a scientific discipline often presage its imminent senescence. »¹⁰⁰ (Stent, 1968)

1.1. Défis et limites de l'approche bibliographique

L'émergence et la diffusion d'une spécialité scientifique sont des phénomènes difficiles à anticiper et à saisir. Les appréhender suppose une compréhension des principes de formation, de circulation et d'évolution des idées et des connaissances d'une part, de constitution, de développement et d'organisation des collectifs scientifiques, d'autre part. Parmi les études sur la science (*science studies*), les études de cas sur les spécialités scientifiques puisent leurs origines dans une longue tradition de recherche. Elles découlent notamment des réflexions très anciennes en philosophie et en histoire des sciences sur le progrès et l'évolution des sciences. Au cours des années 1970, un très grand nombre d'études consacrées aux logiques sociales d'organisation et de développement de spécialités scientifiques furent publiées. Ces études, assez hétérogènes, s'inscrivent dans cet héritage, mais aussi dans la lignée plus immédiate de trois ouvrages : dans l'ordre, en 1962, celui de Thomas Kuhn sur la structure des révolutions scientifiques, la même année, celui de Derek de Solla Price sur la croissance de la science, et en 1972, celui de Diana Crane sur les collègues invisibles (Kuhn, 1962 ; Price, 1963 ; Crane, 1972). Dans les années 1980, c'est avec un peu de recul que les sociologues reviennent sur cette famille de travaux (Chubin, 1985).

Pour être bref, on peut retenir trois idées fondamentales explorées par cette littérature. Premièrement, la progression de la science n'est pas un processus linéaire. La progression scientifique peut-être saccadée, rythmée par des événements, comme l'adoption de nouvelles théories, dont la possibilité, la formulation et l'acceptation sont largement soumis au contexte social et cognitif dans lequel les chercheurs évoluent (voir les développements du Chapitre 1). Deuxièmement, la scientométrie ou l'étude de tout ce qui, en rapport avec la science institutionnalisée, peut être mesuré est un moyen probant de suivre les rythmes de production et d'extraire une image de l'organisation de la production scientifique par l'intermédiaire des relations que l'on peut établir entre les travaux scientifiques (auteurs en commun, institutions en commun, références en commun) (voir

¹⁰⁰ Traduction : « Commencer à écrire l'histoire d'une discipline scientifique est souvent le signe que sa fin est proche. »

Chapitre 5). Troisièmement, la spécialisation scientifique est une nécessité qui s'explique par la croissance de la science — à ce sujet, on peut aussi ajouter la contribution de Ben-David datant de 1966 sur le développement de la psychologie (Ben-David, 1991). Ainsi, le niveau de la spécialité scientifique est un niveau intéressant pour comprendre l'organisation du travail des scientifiques au sein de « collectifs » ou « communautés » et sa traduction en énoncés scientifiques — à ce sujet la contribution de Warren Hagstrom est également culte (Hagstrom, 1965 ; Granjou & Peerbaye, 2011).

Mais après la multiplication des études de spécialités dans les années 1970, il semble que l'engouement soit un peu retombé dans les années 1980-1990. Cette baisse d'intérêt temporaire peut se comprendre par l'intérêt croissant des sociologues des sciences à partir des années 1980 pour les pratiques de laboratoire (Knorr-Cetina, 1982). À partir de ce moment : « Parler de communautés comme d'une évidence, comme d'un *explanans*, ne suffit plus : les regroupements proposés comme pertinents au regard du travail scientifique doivent désormais être empiriquement traçables au regard des engagements et des raisonnements pratiques des acteurs impliqués, et être porteurs de significations pour ces derniers. » (Granjou et Peerbaye, *ibid.*). Le retour des travaux de sociologie sur les spécialités scientifiques se produit dans les années 2000. Parmi les facteurs susceptibles de l'expliquer, outre les questions d'accessibilité des données et les nouvelles possibilités en matière de traitement et de représentation numérique, on note le souci d'établir un lien entre la sociologie des pratiques et connaissances scientifiques et les études sociales de spécialités scientifiques (Zuckerman, 1988 ; Wray, 2005), qui tient notamment à l'impression que les changements conceptuels ont été négligés au bénéfice des facteurs sociaux dans les études de spécialités scientifiques parues dans les années 1970 (Wray, *ibid.*). Il existe aussi des raisons stratégiques liées à l'intérêt des administrateurs de la science pour des méthodes bibliométriques permettant objectivement de repérer le développement de nouveaux fronts de recherche. Enfin, sur un tout autre plan, ce retour peut s'expliquer par la volonté de mieux articuler l'approche historique, géographique et sociale des sciences pour saisir ce type d'objets (Chapitre 4). En particulier, on peut remarquer que dans la plupart des études connues de spécialités, l'espace géographique est plus contextuel que problématisé ou explicatif. Les modèles classiques d'émergence prévoient qu'une spécialité puisse se développer à différents endroits en même temps sans que les chercheurs se connaissent nécessairement ou aient au départ conscience de travailler sur les mêmes problèmes (Mullins, 1962 ; Lemaine, MacLeod, Mulkay, & Weingart, 1976). Mais, à notre connaissance, la question de savoir comment les différentes équipes se mettent à interagir à l'échelle mondiale et le cas échéant à développer le sentiment d'appartenir à une « communauté » malgré la distance et les frontières n'a jamais été clairement posée sous un angle géographique dans le cadre des études de spécialité. À cet

égard, notre démarche s'inscrit dans le mouvement de renouveau des études sur les spécialités scientifiques.

Comme en témoigne la recension très fournie de Steven A. Morris et Betsy Van der Veer Martens sur les études de spécialités, il existe un grand nombre de moyens d'aborder une spécialité (Morris & Van der Veer Martens, 2008). Selon l'angle qui est privilégié, le recours aux données bibliographiques n'est pas même indispensable. Si par exemple on s'intéresse au rôle de l'instrumentation ou des échanges informels entre membres d'un collectif scientifique ou si on se limite à une zone géographique précise, il est sans doute possible de ne pas se préoccuper de constituer et d'analyser un corpus bibliographique (voir par exemple les études de Law, 1973 ; Mounier-Kuhn, 2014). Mais construire un tel corpus est un moyen intéressant de circonscrire et définir une spécialité. Ce travail de définition est au fondement de l'approche bibliographique des spécialités scientifiques.

Cette approche consiste à sélectionner un corpus de documents que l'on juge représentatif du phénomène scientifique étudié (tous les articles de cet auteur, tous les articles publiés dans telle revue ou telle institution, tous les articles qui traitent de telle question ou qui relèvent de telle discipline, tous les articles cités par tel article, tous les articles auxquels fait référence tel article etc.). Plusieurs obstacles justifient néanmoins que l'on adopte une certaine réserve envers cette démarche :

- d'une part, sélectionner revient à délimiter mais les bases de données bibliographiques étant préalablement bornées, le résultat des requêtes est contraint par l'étendue des bases, autrement dit par les revues qui y sont indexées. De ce fait, même en ayant très bien défini les limites de l'objet auquel on s'intéresse, il n'est pas certain que le corpus obtenu parvienne à le résumer de façon satisfaisante ;
- d'autre part, il faut garder à l'esprit que le corpus ne peut faire émerger que la partie « publique » de l'activité des chercheurs (l'expression est de Béatrice Milard). Les motivations des chercheurs ne figurent pas dans leurs publications et sont difficiles à saisir avec une approche seulement quantitative. En particulier, il est généralement admis que ce qui conduit un chercheur à citer telle ou telle référence, à utiliser tel ou tel mot-clef, ou à co-signer un article avec tel ou tel confrère, dépend de sa position dans le « champ scientifique » et du contexte spatio-temporel dans lequel il se trouve (Leydesdorff, 1997 ; Milard, 2010). Donc, en plus de varier dans le temps et l'espace, le statut de la citation et celui de la co-signature n'est pas le même d'une discipline à une autre, ni même d'un chercheur à l'autre. De plus, on ne cite toujours qu'une partie de ses influences (MacRoberts et MacRoberts, 1986 ; Zuckerman, *op. cit.*) et l'on ne co-signe toujours qu'avec une partie de son entourage scientifique (Edge & Mulkay, 1976 ; Stokes & Hartley, 1989).

C'est pourquoi, nous croyons fermement que les données bibliographiques ne sont suffisantes pour un travail à prétention sociologique qu'à condition d'être enrichies de matériaux extérieurs comme les avis ou témoignages d'experts. En effet, les méthodes mixtes ou hybrides enrichissent l'analyse des corpus (Lievrouw, Rogers, Lowe, & Nadel, 1987 ; Laurens, Zitt, & Bassecoulard, 2010). Dès la fin des années 1980, Loet Leydesdorff plaide pour soigner les interprétations en identifiant « un besoin croissant d'intégrer les théories qualitatives issues de la philosophie, la sociologie ou l'histoire des sciences aux perspectives quantitatives propres aux études scientométriques. »¹⁰¹ (Leydesdorff, 1989). Qu'il s'agisse de réaliser une étude bibliométrique sur un auteur, un laboratoire ou un front de recherche, le travail d'élaboration de la requête et d'interprétation des résultats paraît indissociable de l'acquisition de connaissances de fond sur l'objet d'étude. La valeur de l'étude (tant dans sa conception que dans ses résultats) est considérablement améliorée si elle est complétée par un éclairage qualitatif. La connaissance du domaine peut s'appuyer sur une pratique personnelle, sur des avis d'experts ou encore sur des lectures ou des observations. Aucune de ces voies n'est à préférer aux autres. Dans les deux premiers cas, on peut suspecter une trop forte implication (de soi ou des experts) dans la spécialité étudiée, et donc un manque d'objectivité¹⁰². Et si l'on choisit de se construire par soi-même une connaissance du sujet, il est alors difficile d'échapper aux erreurs d'interprétation.

Ces considérations sont importantes notamment car les corpus bibliographiques varient suivant la méthode retenue pour les constituer. Notamment, il a été prouvé qu'il existe une différence significative entre un corpus sélectionné avec avis d'experts et un corpus obtenu à partir d'une analyse de citation (Hurt, 1983). C'est le moment de l'argumentation où il faut convenir qu'il n'y a pas de protocole idéal pour élaborer un corpus de littérature scientifique qui soit parfaitement représentatif d'une spécialité de recherche. En dépit des limites évoquées, nous pensons que les données bibliographiques sont un support de travail riche pour saisir la dynamique, l'organisation sociale et spatiale d'une spécialité scientifique.

¹⁰¹ L'original : « There is a growing recognition of the need to integrate qualitative theorizing in the philosophy, sociology and history of science with the quantitative perspective provided by scientometric studies. » (Leydesdorff, 1989).

¹⁰² Le biais qui consiste à accorder une trop grande confiance aux témoignages recueillis est au cœur des plaidoyers de Michael Mulkay et Nigel Gilbert pour le développement de l'analyse de discours des scientifiques. En 1984, Steven Shapin leur répond au nom des historiens : il est d'après lui possible de reconstituer rigoureusement une suite d'événements passés, donc de faire de l'histoire, à partir de témoignages (Shapin, 1984).

Dans le cas de la réparation de l'ADN, l'approche bibliographique apparaît justifiée par au moins trois raisons :

- Les bases de données bibliographiques présentent une bonne couverture des travaux en biomédecine si bien qu'aux dires de notre informatrice privilégiée (membre de cette spécialité), aucune revue importante ne manque dans les résultats retournés.

- Les chercheurs en biomédecine, en tout cas ceux de la branche étudiée, sont attachés à la visibilité de leurs travaux, et font en sorte de publier dans des revues bien référencées en anglais.

- Les stratégies de publication ne sont pas neutres dans l'affirmation de ce collectif de recherche puisqu'une revue bien référencée a pris le nom officiel de la communauté dans les années 2000 : « *DNA Repair* » ; et qu'on le verra, la référence au mot-clef « *DNA repair* » n'a cessé de progresser depuis qu'il a été retenu comme descripteur en 1971.

La création d'une revue est un des éléments permettant d'affirmer que la réparation de l'ADN est devenue une « spécialité » et, au-delà, que les membres de cette spécialité font ou, du moins certains d'entre eux, entendent faire « communauté ». En effet, tout comme la scientométrie s'est affirmée comme spécialité à part entière en 1978 avec la création d'une revue appelée logiquement *Scientometrics*, le collectif qui nous intéresse dispose de sa revue aujourd'hui appelée *DNA Repair*. Celle-ci fait partie d'une collection de revues ayant pour ancêtre commun la revue *Mutation Research* fondée en 1962 aux Pays-Bas (Natarajan, 2007). Mais le fait qu'il ait fallu attendre les années 2000 pour que *DNA Repair* devienne une revue autonome suggère que la réparation de l'ADN a mis du temps pour accéder au rang de spécialité scientifique en biologie moléculaire au même titre que la transcription de l'ADN ou la réplication de l'ADN.

L'existence de cette revue indique qu'il existe une volonté d'institutionnalisation du collectif et de distinction à l'égard d'autres groupes de recherche, mais, comme on le verra plus loin, cela ne signifie pas que tous les chercheurs qui pourraient s'y reconnaître partagent cette volonté. Il est intéressant de noter que la revue dispose d'une section « *Historical reflections* » participant à forger le récit des origines particulier à tous collectifs de type communautaire. Errol C. Friedberg y participe volontiers : il entretient, depuis le Texas, une passion pour l'histoire à côté de son rôle de rédacteur en chef de la revue et son statut de pionnier dans le champ des recherches sur les mécanismes de réparation de l'ADN. Son ouvrage *Correcting the Blueprint of Life : An Historical Account of the Discovery of DNA Repair Mechanisms* en témoigne, qui est précédé par la dédicace suivante : « À mes professeurs d'école, qui m'ont donné le goût de l'histoire et à la communauté internatio-

nale de la réparation de l'ADN, dont la convivialité et le soutien m'entourent depuis un si grand nombre d'années. »¹⁰³ (Friedberg, 1997).

Cet ouvrage insiste justement sur l'âge de la communauté pour mettre en avant le temps écoulé avant qu'elle accède à la reconnaissance académique. Ce thème est prégnant parmi les membres du collectif et on le retrouve par exemple dans le récit de Peter Strike, professeur à Édimbourg, paru sous le titre « *Bacteria and DNA repair – 50 years together* » : « La réparation de l'ADN dispose maintenant d'une place au centre, à côté de la réplication de l'ADN et du cycle cellulaire, comme processus essentiel au fonctionnement de chaque cellule. Il est bien sûr imprécis de la décrire comme un seul phénomène puisqu'il s'agit d'un assemblage de fonctions, capables de réagir à presque tous les défis que l'environnement impose à notre ADN, en préservant la structure moléculaire et l'intégrité génétique. Les déficits en termes de réparation ont des effets sur la survie, les mutations, le cancer, la fabrication d'anticorps, la reproduction, la méiose, la mitose... la liste est infinie. Et pourtant, il n'y a pas si longtemps, la réparation de l'ADN était encore considérée comme un curieux petit domaine en biologie, essentiellement étudié de manière peu conventionnelle par des "radiobiologistes". Le sujet était souvent en dehors des intérêts dominants, et effectivement, sa légitimité était très limitée avant que ne soit découvert un lien entre la réparation des mésappariements et le cancer du côlon en 1980. Presque aucune analyse biochimique des fonctions de réparation n'avait été tentée avant les années 1960 quand bien même de bons généticiens avaient mis en évidence la présence de mécanismes réparateurs bien avant cela. »¹⁰⁴ (Strike, 2003).

La dernière phrase de la citation est intéressante parce qu'elle suggère que les premiers jalons permettant le développement de la spécialité scientifique ont été posés avant les années 1960. C'est aussi un élément sur lequel insiste Friedberg dans le premier chapitre de son livre qu'il conclut comme suit : « En bref, la période allant du milieu des années 1930 à la fin des années 1940 peut être vue comme celle de l'émergence du do-

¹⁰³ L'original : « To my schoolteachers, who gave me a love of history and the international DNA repair community, whose collegiality and support I have enjoyed for so many years. » (Friedberg, 1997).

¹⁰⁴ L'original : « DNA repair now sits centre stage, along with DNA replication and the cell cycle, as one of those key processes essential to the function of every living cell. It is of course inaccurate to describe it as a single process – it is a comprehensive suite of functions, able to deal with almost every environmental challenge thrown at our DNA, maintaining the molecule's structure and genetic integrity. Defects in repair impact on survival, mutation, cancer, antibody production, reproduction, meiosis, mitosis...the list is endless. And yet, not too long ago, DNA Repair was considered an odd little corner of biology, largely studied in a rather oblique manner by "radiobiologists". The subject was generally outside mainstream interests, and indeed it could be considered that a general realization of the importance of repair processes really only came with the demonstration of the unequivocal link between mismatch repair and colon cancer in the 1980s. Almost no biochemical analysis of repair functions was even attempted until the 1960s, although good genetics had begun to indicate the presence of these processes somewhat earlier. » (Strike, 2003).

maine de la réparation de l'ADN et de la mutagenèse, mais aussi plus généralement de tout ce qui est aujourd'hui compris dans la biologie moléculaire comme discipline (...) Mais un ensemble de biais intellectuels, et dans une moindre mesure, de facteurs politiques, ont retardé l'émergence de la réparation de l'ADN en tant qu'aire de recherche légitime. On pensait que les gènes étaient faits de protéines et on les considérait comme intrinsèquement stables (...) On imaginait que les mutations étaient de rares événements ayant un intérêt seulement pragmatique pour les études génétiques, mais on croyait que leurs mécanismes et leurs origines n'étaient pas traçables. La guérison après l'exposition aux rayons X et aux UV n'était rien de plus qu'un phénomène anecdotique, voire dans le pire des cas, une préoccupation des scientifiques du gouvernement dont la fonction était de trouver des applications biologiques à l'utilisation militaire des radiations, une tâche pour laquelle ils étaient largement subventionnés. Donc, la première preuve directe de réparation de l'ADN n'a pas été avancée avant le milieu du siècle, et il a fallu encore une décennie de plus pour que le terme soit clairement intégré au lexique de la biologie moléculaire et cellulaire. »¹⁰⁵ (Friedberg, 1997, p. 24-25).

D'après l'étude de l'historien des sciences Doogab Yi sur cette question, le retard peut effectivement s'expliquer par le caractère contre-intuitif de ce progrès scientifique : l'idée que l'ADN puisse s'endommager allait à l'encontre de la croyance dans la stabilité des gènes et dans l'irréversibilité des dommages causés par les radiations sur le corps humain (Yi, 2007). Au-delà, Yi insiste sur le rôle essentiel que les biophysiciens et radiobiologistes de Yale ont joué dans l'abolition de ces croyances. C'est à la fois parce qu'ils étaient convaincus par l'intérêt médical des recherches sur les radiations et parce qu'ils ont su décrire leurs observations à l'aide du vocabulaire de la théorie de l'information (en parlant par exemple de « corrections d'erreurs ») que le petit groupe de Yale serait arrivé à se distinguer. Le choix de ce registre de langage a permis, d'après Yi, d'étendre l'influence des découvertes en dehors du champ de la radiobiologie et de susciter l'intérêt des biochimistes. Comme Nicholas Mullins l'a montré à propos du Groupe Phage, la « théorie de

¹⁰⁵ L'original : « In brief, the period from the mid nineteen-thirties to the late nineteen-forties may be characterized as one during which not only the field of DNA repair and mutagenesis emerged, but also much of what we now designate as the discipline of molecular biology [...] But a combination of intellectual biases, and to a lesser extent, political influences, constrained the emergence of gene/DNA repair as an area of investigative inquiry in parallel with other aspects of gene function. Genes were presumed to be made of proteins and to be intrinsically stable [...] Mutations were considered to be rare event that were of enormous pragmatic value for genetic studies, but their mechanism or origin was not obviously experimentally tractable. Recovery after exposure to X rays and UV light was an anecdotal phenomenon at best, and at worst the province of government scientists who were primarily intent on gleaning useful biological applications for the militaristic use of radiation, a task for which they were lavishly supported. So, the first direct experimental evidence for DNA repair did not emerge until just before the middle of this century, and it was not until almost a decade later that the term was confidently and unambiguously incorporated into the lexicon of molecular and cellular biology. » (Friedberg, 1997, p. 24-25).

l'information » a joué un rôle important dans le développement et l'autonomisation de la biologie moléculaire (Mullins, *op. cit.*). Il aurait donc fallu que les pionniers de la réparation de l'ADN adoptent eux-aussi le vocabulaire de l'information pour que les tenants du Groupe Phage (organisés autour de Max Delbrück et surtout préoccupés par les fonctions de réplication) acceptent l'idée que l'étude des fonctions de réparation puisse également relever de la biologie moléculaire.

Cette transition correspond au moment où la terminologie « réparation de l'ADN » est adoptée. Avant les années 1960, bien qu'un premier mécanisme de réparation ait déjà été découvert (la photoréactivation) et malgré les indices suggérant l'existence d'autres mécanismes, ne nécessitant pas pour leur part l'intervention de la lumière, le terme n'est presque jamais utilisé. Les mots « guérison » et « survie » sont employés indifféremment là où on parlerait aujourd'hui de « réparation ». À présent, il existe plus de 70 000 documents indexés avec « *DNA repair* » dans le titre (mais aussi, quand ils sont renseignés, dans le résumé et la liste de mots-clefs des publications). Les premiers datent du milieu des années 1960 et sont effectivement le fruit des chercheurs passés par le groupe de Yale, on y reviendra. Ensuite, la plupart des pionniers ont publié leur premier article « *DNA repair* » entre 1970 et 1973, au moment où « *DNA repair* » fait son entrée dans le thésaurus de la *National Library of Medicine* américaine (Friedberg, 1997). Aussitôt, le nombre de publications spécifiant « *DNA repair* » décolle. Dans *Scopus*, il passe de 8 publications en 1969, à 139 en 1971 puis 565 en 1975. Dès lors, plus que la formation du collectif, les corpus délimités à partir de ce mot-clef permettent de suivre la diffusion de la formulation depuis les États-Unis. En effet, la localisation des premiers articles spécifiant « *DNA repair* » ne coïncide pas forcément avec la date d'apparition de la thématique de recherche dans ces lieux. À l'époque où l'utilisation du mot-clef n'est pas encore stabilisée (entre 1965 et 1975), il serait donc vain de se focaliser sur la chronologie annuelle de son adoption. En revanche, nous entendons comparer la carte de la répartition du mot-clef au stade d'émergence (1960-1975) et au stade de la maturité (2006-2008) afin d'évaluer l'ampleur et les logiques géographiques de son expansion. Au-delà, la carte des recherches au stade d'émergence permet de repérer les foyers de la spécialité.

1.2. La géographie de la réparation de l'ADN : quelle évolution ?

Comme « *DNA repair* » est devenu le terme officiel pour identifier les travaux des chercheurs intéressés par l'étude des mécanismes de réparation de l'ADN dès le début des années 1970, il apparaît justifié de mobiliser ce simple descripteur pour délimiter un corpus bibliographique représentatif du long processus ayant conduit la réparation de l'ADN à prétendre au rang de spécialité en biologie moléculaire dans les années 2000.

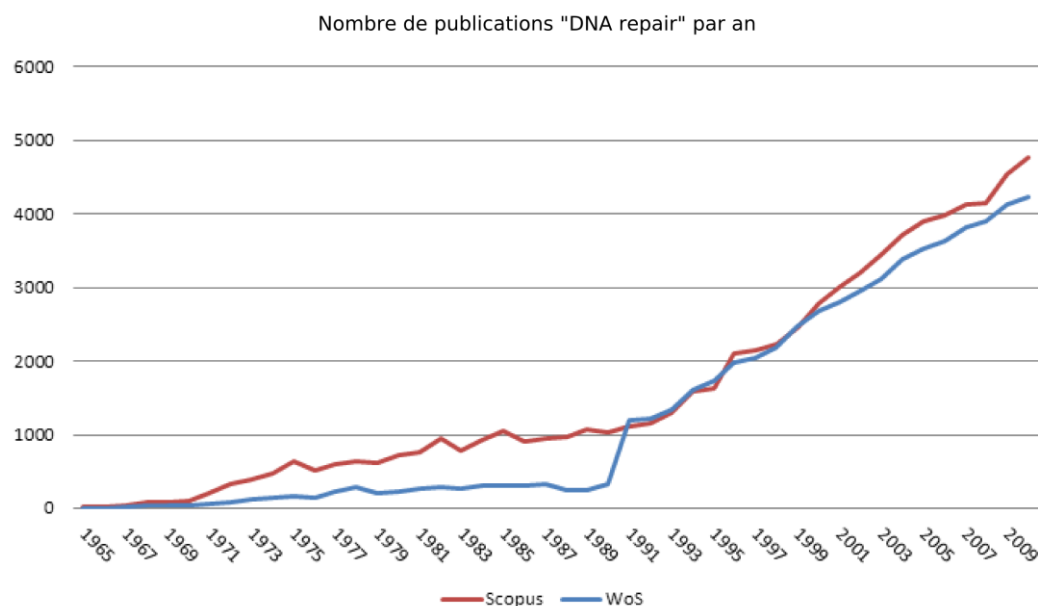


Figure 70 – Nombre de publications « DNA repair » par an et par source bibliographique

D'après la Figure 70, le nombre de documents indexés par le mot-clef « DNA repair » a commencé à augmenter au début des années 1970 et a continué de progresser jusqu'à nos jours. Toutefois, entre 1960 et 1990, l'augmentation est moins marquée dans le WoS que dans *Scopus*. Après 1990, l'évolution du nombre de publications par an est en revanche comparable entre les deux sources. Aux alentours de 1990, on repère un pic dans la couverture du WoS qui s'explique par le référencement des résumés et des mots-clefs associés à chaque publication, alors que seuls les titres ont été correctement indexés dans la base pour les documents parus avant cette période. Aussi, pour les documents indexés avant 1990, les adresses des auteurs sont beaucoup plus souvent renseignées dans les enregistrements de *Scopus* que dans ceux du WoS. Or, l'information contenue dans les adresses est indispensable pour saisir la géographie du domaine de recherche. Cette différence entre le WoS et *Scopus* vient du fait que, pour les années 1960-1990, *Scopus* rassemble des notices bibliographiques issues de plusieurs bases ce qui, en contrepartie, entrave la significativité du résultat. Numériquement, le résultat de la requête effectuée dans *Scopus* est d'ailleurs à peu près équivalent à celui obtenu lorsqu'on formule la requête dans le *Web of Knowledge* (WoK), qui est un ensemble hétérogène de bases de données comprenant le WoS. Mais, en raison du plus grand nombre de renseignements disponibles dans *Scopus*, nous avons privilégié *Scopus* pour décrire la géographie de la réparation de l'ADN au stade d'émergence de la spécialité (1965-1975).

Y compris dans *Scopus*, la qualité de l'indexation des documents parus entre 1965-1975 est somme toute assez mauvaise et certaines adresses ne sont pas enregistrées du tout. En particulier, en cas de collaboration entre plusieurs laboratoires, il est fréquent qu'une seule adresse soit spécifiée ce qui empêche de caractériser ces collaborations. Pour pallier ce problème, quand aucune adresse n'était renseignée pour un article donné, nous avons complété par l'affiliation d'au moins un des auteurs de l'article (lorsqu'elle était disponible dans le corpus). En itérant ainsi les affiliations d'auteurs pour réduire les données manquantes, cette méthode est évidemment responsable de l'introduction d'approximations que nous avons jugées préférables à l'absence d'information. En suivant cette procédure, seules 129 publications (sur 1804 articles, recensions et lettres recensés entre 1965 et 1975) n'ont pas pu être localisées.

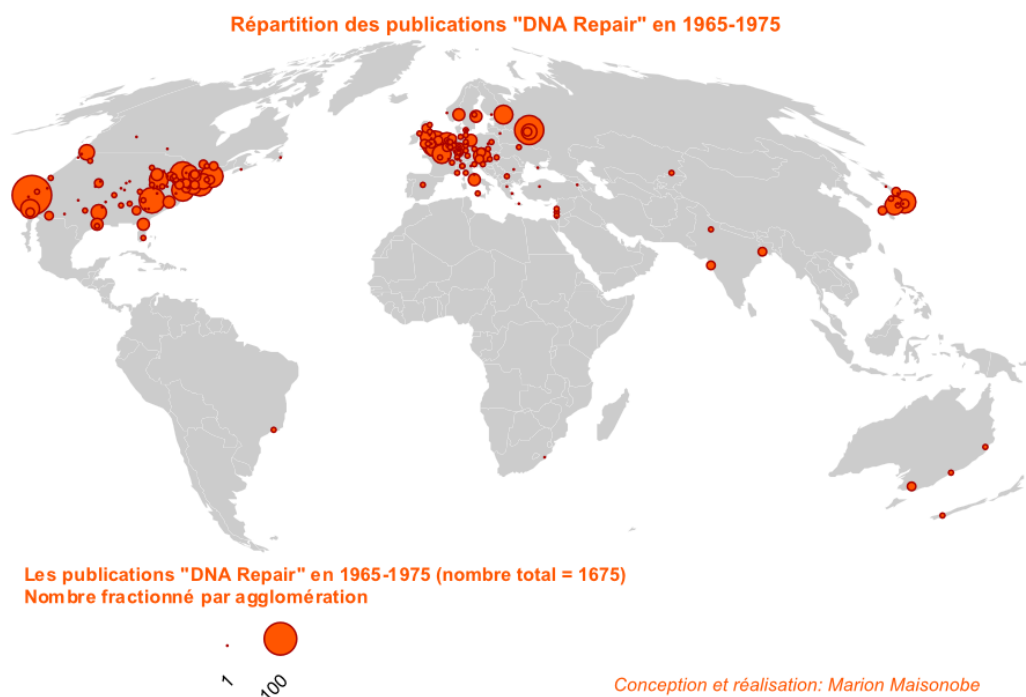


Figure 71 – Répartition des 1675 publications avec le mot-clef « DNA Repair » (1960-1975). Source : Scopus

Ensuite, les adresses associées aux 1675 publications restantes ont été « nettoyées » et harmonisées pour ne conserver qu'une variante possible de dénomination par adresse de laboratoire. Après l'harmonisation, un champ « ville » a été isolé à partir du champ « adresse » pour être mis en relation avec la table de correspondance « ville » - « agglomérations urbaines » conçue dans le cadre du programme ANR « Géoscience » (voir Chapitre 6). Les données de cette opération peuvent alors être résumées dans une carte de la répartition mondiale des publications scientifiques utilisant la terminologie « DNA repair » entre

1965 et 1975. Cette carte permet d'identifier les espaces de concentration de l'activité scientifique sur la réparation de l'ADN et de repérer les foyers de recherche. On en distingue cinq auxquels sont associés des volumes de publication considérables : aux États-Unis, sur la côte californienne et la mégapole de la côte Est ; de l'autre côté de l'Atlantique, en Europe du Nord-Ouest, en Russie et au Japon (Figure 71).

Les quantités de publications par agglomération représentées sur la carte ont été calculées à l'aide d'un comptage entier fractionné (*Whole Normalized Counting*). Pour un article donné, ceci implique que le crédit accordé à une ville donnée dépend du nombre de villes impliquées dans la signature de l'article (Partie 2.Chapitre 7). Mais, puisque très peu de collaborations scientifiques sont correctement renseignées dans ce corpus et que le nombre moyen d'auteurs par article est beaucoup moins important à l'époque qu'aujourd'hui, le choix de ce mode de comptage n'a pas d'incidence particulière sur les résultats.

En revanche, ce choix a été décisif lorsqu'il a été appliqué au corpus de publications « *DNA repair* » sélectionné pour rendre compte de la géographie contemporaine (2006-2008). Il présente de plus l'avantage de normaliser les ordres de grandeur entre les deux périodes car malgré l'augmentation des co-signatures intervenue entre temps, l'unité de compte demeure l'agglomération et non l'adresse. Néanmoins, comme les pratiques de publications ont beaucoup changé (le nombre d'adresses par article notamment) ainsi que la qualité du référencement des publications (en particulier des articles co-signés), nous préférons éviter de comparer directement les valeurs brutes des deux corpus. C'est pourquoi, dans ce qui va suivre, la comparaison entre les deux périodes s'appuie d'une part sur les données indiquant la présence ou l'absence d'activité (quels sont les lieux d'où a été signé au moins une publication « *DNA repair* » ?), et d'autre part sur les indices relatifs de dispersion/concentration permettant de traiter chaque distribution sur un pied d'égalité.

Pour obtenir une représentation de la géographie actuelle du domaine, nous considérons l'ensemble des publications indexées entre 2006 et 2008 dans le WoS avec pour mot-clef : « *DNA repair* ». Comme le WoS et *Scopus* donnent des résultats équivalents pour cette période, il n'y a pas de raison de préférer une source à l'autre. Le niveau d'activité des années 2000 étant au moins 4 fois supérieur à celui du stade d'émergence (Figure 72), une fenêtre temporelle de 3 ans est suffisante pour saisir la géographie contemporaine. Ce faisant, 5086 publications parues entre 2006 et 2008 ont été extraites du WoS dont la répartition par agglomération peut être mise en regard avec celle des 1675 articles parus entre 1965 et 1975 avec le mot-clef « *DNA repair* » (Figure 71 et Figure 72).

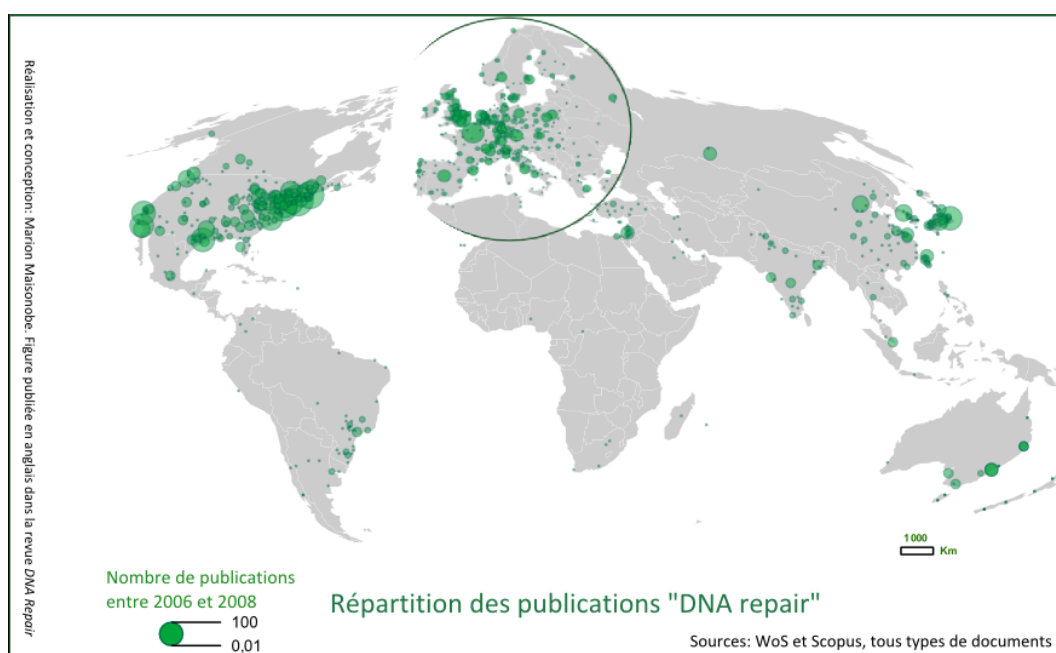


Figure 72 – La répartition géographique des publications « DNA repair » en 2006-2008.
Source : Web of Science

Tant que l'on se limite à analyser la présence ou non d'activité (quels sont les lieux d'où a été signé au moins une publication « DNA repair » ?), la comparaison montre que 800 agglomérations sont actives sur le sujet en 2006-2008, soit 4 fois plus qu'au stade d'émergence (195). Sur les 195 lieux d'activité identifiés à l'époque, 175 sont toujours actifs entre 2006 et 2008 (en rouge sur la Figure 73).

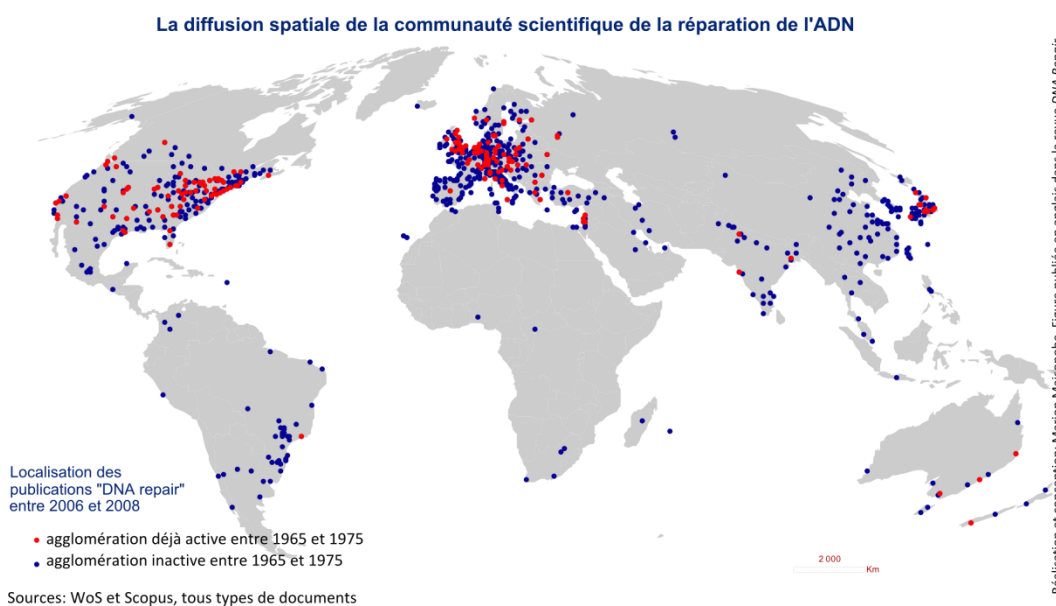


Figure 73 – Lieux d'où sont publiés des documents avec le mot-clef « DNA repair » en 2006, 2007 et 2008. Source : Scopus et Web of Science

Parmi les 175 « agglomérations pionnières » toujours actives, 60 sont aux États-Unis et le reste se trouve principalement en Europe, au Japon et au Canada. Bien qu'une grande part de la diffusion se soit effectuée au voisinage des agglomérations pionnières, la carte suggère que la communauté s'est étendue à l'échelle mondiale jusqu'à atteindre des pays qui n'étaient pas impliqués au départ. Ces deux observations ne sont pas contradictoires : 77 % des « nouveaux » lieux se trouvent dans des pays déjà actifs au stade d'émergence ; tandis que 11 « nouveaux » pays comptent plus de 4 agglomérations actives en 2006-2008 : La Chine, la Corée du Sud, le Mexique, le Portugal, Taiwan, l'Argentine, l'Irlande, l'Iran, l'Afrique du Sud, la Thaïlande, et la Tunisie. La Chine est le pays qui a enregistré la croissance la plus spectaculaire. Alors qu'elle a seulement intégré le domaine de recherche dans les années 1980, elle abrite aujourd'hui une quarantaine d'agglomérations actives. Seulement deux pays disposent d'autant de lieux d'activité : l'Italie avec 43 et les États-Unis avec 126. Venant après, il y a l'Espagne, la France, le Brésil et le Japon au sein desquels le processus de diffusion a été le plus remarquable.

Ce phénomène de diffusion témoigne bien du fait que le domaine de la réparation de l'ADN s'est développé entre les deux périodes. Ce développement a été favorisé par les relations de plus en plus nombreuses qui ont été établies entre les mécanismes de réparation de l'ADN, ou plus exactement leurs défaillances pouvant provoquer des mutations (sinon la mort cellulaire), et des enjeux essentiels comme le cancer et le vieillissement. Toutefois, on considère parfois qu'il est plus difficile d'intégrer un domaine de recherche longtemps après son émergence. Certains décrivent ce processus comme une course : les derniers à partir auraient plus de mal à rattraper les premiers. Au niveau des lieux, ce phénomène est parfois qualifié de « *place dependence* » (expression inspirée de l'idée de dépendance du sentier) qui s'explique par l'héritage laissé par les primo-adoptants en un lieu donné (Arthur, 1988 ; Boschma & Frenken, 2006). Afin de vérifier cette représentation, il convient d'analyser plus finement l'évolution du poids des agglomérations pionnières dans la production globale.

1.3. La géographie de la réparation de l'ADN : quel poids du passé ?

Entre 2006 et 2008, la production de publications portant sur la réparation de l'ADN est moins concentrée qu'elle ne l'était au stade d'émergence. En 1965-1975, les 10 premières agglomérations les plus publiantes produisent 40 % de la production alors qu'elles ne participent plus qu'à 23 % de la production totale en 2006-2008. Pour autant, comme le suggèrent les cartes précédentes de répartition de la production (Figure 71 et Figure 72), les lieux actifs au stade d'émergence continuent de concentrer une grande part de la production entre 2006 et 2008. Prises simultanément, les 175 agglomérations pion-

nières sont responsables de 65 % de la production en 2006-2008 (3304 publications sur 5086), et on trouve parmi ces 175 agglomérations l'essentiel des lieux les plus dynamiques en 2006-2008. Ceci étant, les agglomérations les plus publiantes en 2006-2008 ne sont pas forcément celles qui réalisaient le plus de publications au stade d'émergence. Autrement dit, les espaces pionniers sont toujours en activité au stade de la maturité mais la hiérarchie entre ces espaces s'est partiellement transformée (Figure 74 et Figure 75).

Evolution de la répartition des publications (en part relative) entre agglomérations

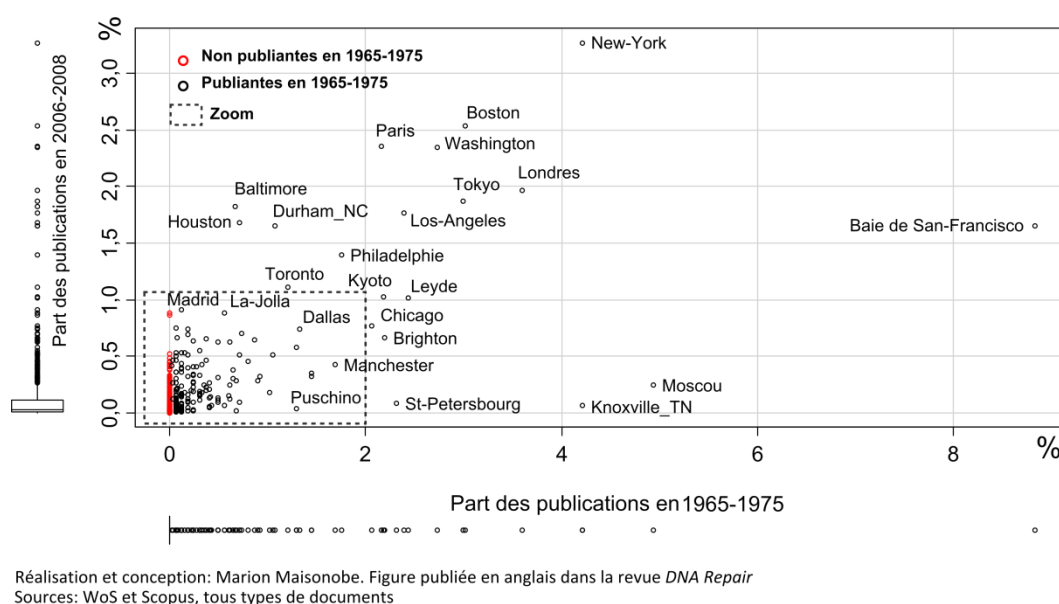
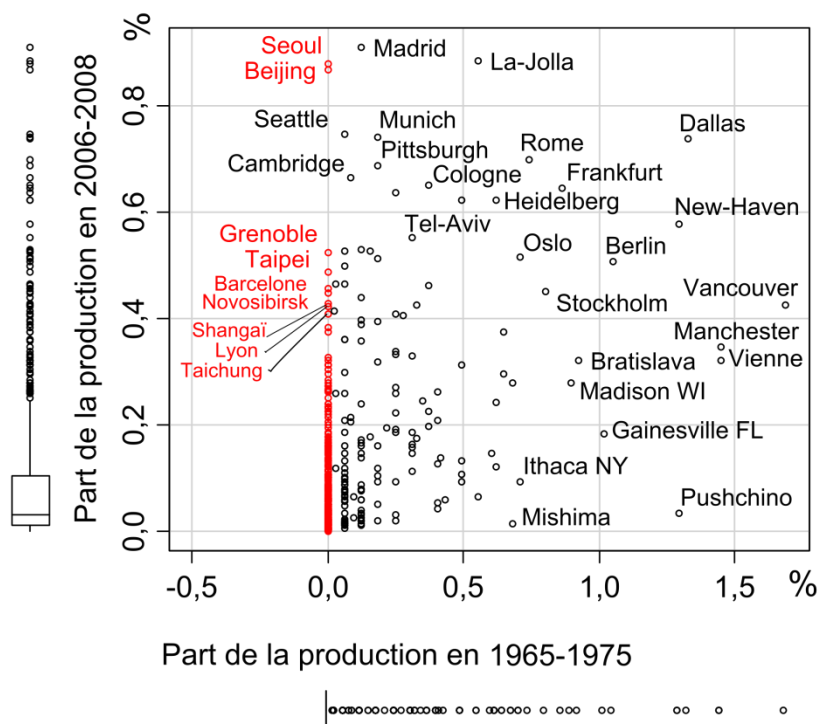


Figure 74 – L'évolution de la part de publications avec le mot-clef « DNA repair ». Source : Scopus et Web of Science

Alors que la baie de San-Francisco concentrait 9 % de la production en 1965-1975 (143 publications, rang n°1), seulement 1,5 % de la production en provient en 2006-2008 (84 publications, rang n°11). Ainsi, la côte Est des États-Unis est passée devant la Californie entre les deux périodes. En particulier, à la troisième place en 1965-1975, New-York se retrouve première en 2006-2008 (avec 166 publications, 3 % du total). Contrairement à ces zones dont l'activité est restée au top depuis le début, Moscou et Knoxville (TN), respectivement à la 2ème et 3ème place (*ex-aequo* avec New-York : 68 articles) au stade d'émergence (chacun contribuant à plus de 4 % de la production totale) n'ont pratiquement pas participé à la production en 2006-2008 (poids inférieur à 1 %) (Figure 74 et Figure 75). Ces deux cas correspondent aux cas les plus extrêmes de déclin identifiés dans l'analyse. Ils méritent d'autant plus qu'on s'y arrête qu'ils sont chacun représentatifs d'une cause différente de déclin dont ont été victimes quelques une des équipes ayant travaillé sur la réparation de l'ADN au stade d'émergence. D'une part, le déclin de Moscou

est typique du déclin des équipes soviétiques impliquées dans la « réparation de l'ADN » dès les origines du champ ; et d'autre part, le déclin de Knoxville (Tennessee), abritant les laboratoires fédéraux d'Oak Ridge, est typique des coupes budgétaires intervenues dans les laboratoires fédéraux et le portefeuille militaire de la recherche à partir des années 1970-1980.

Zoom sur l'évolution de la part de publications "DNA repair"



Sources: WoS et Scopus, tous types de documents

Clefs de lecture: Quelle est la participation d'une unité géographique donnée dans la production totale de publications au sujet de la réparation de l'ADN lors de la genèse du domaine (1965-1975) et 35 ans plus tard (2006-2008)? Alors que 5% de la production est réalisée à Moscou en 1965-1975, c'est seulement 0,25% des publications de 2006-2008 qui ont été signées depuis Moscou. Alors qu'elles ne publiaient pas en 1965-1975, Seoul et Beijing ont contribué chacune à plus de 0,8% de la production en 2006-2008.

Figure 75 – Zoom sur l'évolution de la part de publications avec le mot-clef « DNA repair ».

Source : Scopus et Web of Science

Aux laboratoires d'Oak Ridge (Knoxville), le déclin est très précisément lié aux réductions de budget et aux transformations qui ont affectées la division biologie. Cette division a perdu 40 % de ses membres au début des années 1980 et a été incitée à se con-

centrer sur des recherches plus appliquées¹⁰⁶. Son chef, Richard Setlow, ayant commencé sa carrière à Yale, a dû quitter Oak Ridge pour Upton (le laboratoire fédéral de Brookhaven). Après lui, les changements d'affiliations des membres de la division biologie témoignent entre autres du fait que Sankar Mitra a quitté Oak Ridge pour l'Université de Galveston au Texas, qui fait maintenant partie de la « *Texas Mafia* » : le groupe des équipes texanes spécialisées dans la réparation de l'ADN (Friedberg, 1997, p. 50-56). En fait, à partir des années 1970, le rôle des instituts de recherche d'État, largement financés au sortir de la guerre 1939-1945, a été mis en cause dans plusieurs pays (Joyce, 1981). En particulier, l'idée s'est répandue que la recherche fondamentale devait retrouver sa place dans les universités. C'est pourquoi, la plupart des équipes initialement localisées dans des instituts nationaux de recherche nucléaire ont été amenées à déménager. Aux Pays-Bas, les scientifiques du laboratoire national de Rijswijk (une unité du TNO, l'institut national de recherche appliquée) se sont répartis entre l'équipe de Leyde spécialisée dans la recherche sur les mutations cellulaire et l'équipe de biologie cellulaire tout juste créée dans la nouvelle Université de Rotterdam, on y reviendra. Par ailleurs, parmi les quelques endroits pionniers qui ne publient plus du tout sur la réparation de l'ADN en 2006-2008, on compte les centres de recherche nucléaire canadiens de Chalk River et Pinawa. Ce déplacement depuis les centres de recherche nucléaire vers les universités et les centres de recherche sur le cancer et le vieillissement a été un des principaux facteurs de changement pour la géographie de la réparation de l'ADN.

Le second facteur de déclin est lié à l'effondrement de l'URSS. À Moscou, les deux instituts principaux étaient : l'institut de génétique de l'Académie des Sciences dirigé par Nicolay Dubinin (un fameux opposant de Lyssenko) et l'institut d'Épidémiologie et de Microbiologie N.F. Gamelaya de l'Académie de Médecine dirigé par Adelina Skavronskaya. Après une période faste, l'arrêt de leur activité de recherche sur la réparation de l'ADN s'explique par le déclin progressif de la science soviétique et son isolement relatif vis-à-vis des équipes occidentales, on y reviendra. Dans les années 2000, contrairement à ce qu'on aurait pu attendre, c'est à Novossibirsk plutôt qu'à Moscou que l'activité scientifique a repris en Russie alors qu'il n'y avait dans cette ville aucune recherche menée sur la réparation de l'ADN au stade d'émergence. Ainsi, Novossibirsk est l'un des rares endroits qui compte aujourd'hui une équipe spécialisée et remarquablement active dans la réparation de l'ADN, alors qu'aucune activité n'y était recensée en 1965-1975 (en rouge sur la Figure 75).

¹⁰⁶ Source disponible en ligne : Oak Ridge BIENNIAL REPORT 1982-1983. Information évoquée également dans l'historique de l'institution, Source : URL : <http://web.ornl.gov/info/ornlreview/rev25-34/chapter8.shtml>

Arrêtons-nous plus précisément sur ces endroits qui ont réussi à intégrer le champ sur le tard.

Parmi les 20 agglomérations les plus publiantes en 2006-2008, 18 étaient déjà en activité en 1965-1975. Mais, certaines d'entre elles (comme Madrid) étaient parmi les agglomérations les moins impliquées au stade d'émergence. Effectivement, seulement deux publications « *DNA repair* » ont été signées depuis Madrid en 1965-1975. Il s'agissait de la seule ville espagnole active à l'époque. Entre temps, la fin du franquisme et l'intégration de l'Espagne à l'Union Européenne ont permis à la science espagnole de se redynamiser (Grossetti & Losego, 2003). Dans le domaine de la réparation de l'ADN, l'Espagne fait maintenant partie des pays les plus impliqués. Avec un total de 115 publications parues sur le sujet entre 2006 et 2008, l'Espagne est parmi les 10 pays les plus publiants. Par ailleurs, la Chine est le seul des 10 pays actuellement en tête de peloton dont aucune équipe ne travaillait sur la réparation de l'ADN en 1965-1975. Dans le top 10, la Chine a remplacé la Russie. Aussi, les seules agglomérations à avoir intégrées le top 20 alors qu'elles étaient absentes au stade d'émergence sont asiatiques. Il s'agit de Séoul et Beijing. Cette évolution est cohérente avec la tendance générale à l'intégration du monde asiatique dans le système scientifique mondial, tendance observée sur la dernière décennie à partir des données du WoS dans le chapitre précédent (Chapitre 9). Mais, lorsqu'on se penche sur les données de citations, on observe que l'intégration du monde asiatique dans la communauté de la réparation de l'ADN est beaucoup moins évidente. Alors qu'elles sont parmi les 20 agglomérations les plus publiantes sur le sujet en 2006-2008, Séoul et Beijing sont remplacées par Zurich et Oxford parmi les 20 agglomérations les plus citées. Au contraire, les 18 autres agglomérations du top 20 sont à la fois parmi les plus publiantes et les plus citées. Cela suggère que pour les équipes scientifiques entrées tardivement dans le sujet, il existe un décalage ou effet de retard entre le niveau de production et le degré de visibilité. Dans le cas de Séoul et Beijing, ce décalage est considérable puisque Beijing est seulement la 28^{ème} agglomération la plus citée et Séoul, la 35^{ème}.

En fin de compte, le déclin de la science soviétique et le dynamisme scientifique de l'Asie orientale expliquent une grande partie des changements observés. Mais il serait réducteur de considérer l'Asie comme le seul espace ayant participé à l'expansion géographique du domaine de la réparation de l'ADN. Effectivement, l'apparition d'équipes spécialisées sur le tard est un phénomène rencontré dans plusieurs des pays déjà actifs au stade d'émergence. Alors qu'elles sont à proximité d'équipes déjà impliquées au départ, Grenoble et Lyon en France, Barcelone en Espagne, la ville de Québec au Canada, San-Antonio au Texas abritent désormais des équipes qui sont parmi les plus actives du domaine (plus de 20 articles publiés en 2006-2008). Ces cas particuliers suggèrent que les opportunités de « rattrapage » ne se sont pas limitées à l'Asie. La diffusion des recherches

sur la réparation de l'ADN est intervenue simultanément à différentes échelles : à la fois à l'échelle mondiale et à l'échelle de plusieurs pays.

Malgré l'intérêt des éléments de comparaison obtenus à partir des deux corpus et leur adéquation avec les évolutions observées au niveau de l'ensemble des activités scientifiques au cours de la période contemporaine, les conclusions doivent être nuancées au regard de la méthode utilisée. Puisque l'analyse repose sur le recours au mot-clef « *DNA repair* », on pourrait effectivement considérer que les observations ne font que traduire une « banalisation » du terme dans le milieu scientifique. C'est d'autant plus probable que comme l'indique Friedberg, la terminologie a évolué dans le temps et s'est élargie de sorte que : « le champ [de la réparation de l'ADN] désigne désormais plus exactement l'ensemble des réponses biologiques aux lésions d'ADN, et dans ce contexte même le terme "*DNA damage*" a été élargi pour comprendre des phénomènes tels que l'arrêt de la synthèse d'ADN en l'absence de lésions définies. »¹⁰⁷ (Friedberg, 2007).

En même temps, tandis qu'il ressort des récits des pionniers de la réparation de l'ADN qu'il s'agit d'une spécialité résultant de la rencontre de plusieurs branches de recherche (radiobiologie, photobiologie, biophysique, biologie cellulaire, puis moléculaire...), les entretiens avec notre informatrice privilégiée ainsi qu'avec des spécialistes néerlandais de la réparation de l'ADN ont suggéré qu'il existe dorénavant autant de branches internes à la spécialité que de mécanismes de réparation de l'ADN distincts, soit un peu moins d'une dizaine. Aussi, les articles de synthèse, à teneur historique ou pédagogique signalent l'existence de ces nombreuses branches ou sous-spécialités (Figure 76). La variété des organismes, les maladies témoins des déficits de réparation, les différents types de lésions justifient que se soient assez rapidement développés plusieurs groupes de recherche au sein de la communauté. Il est même arrivé que des spécialistes en réparation de l'ADN entrent en interaction avec d'autres groupes de biologistes moléculaires. C'est ce qui s'est produit dans les années 1990, par exemple, entre des spécialistes du mécanisme de réparation par excision de nucléotide (NER) et des spécialistes de la transcription de l'ADN, on y reviendra dans le Chapitre 11.

Dès lors, on peut se demander ce qu'il reste de « la réparation de l'ADN » comme spécialité. En effet, notre informatrice admet volontiers que « la configuration d'ensemble du domaine s'est complexifiée et que très peu de scientifiques disposent maintenant d'une vision d'ensemble de la communauté. »¹⁰⁸ (Maisonobe *et al.*, 2013). Dans les faits, ce type de croissance scientifique par ramification est assez commun. Plus précisément, il est dé-

¹⁰⁷ L'original : « The broader field is now more appropriately referred to as the field of biological responses to DNA damage, and in this context even the term "DNA damage" has been broadened to include phenomena such as the arrest of DNA synthesis in the absence of defined DNA damage. » (Friedberg, 2007).

¹⁰⁸ L'original : « As it grows, the field has become more complex and no scientist can now have a complete overview of this scientific community. » (Maisonobe, Giglia-Mari & Eckert, 2013).

crit et qualifié de « *branching* » par Mulkey et Edge dès les années 1970 à partir de leur étude sur le développement de la radioastronomie en Angleterre (Mulkay, 1975 ; Edge et Mulkay, 1976 ; Mulkay et Edge *in* Lemaire *et al.*, 1976) et on le retrouve décliné par Abbott dans le cadre d'une réflexion qui porte sur le niveau des disciplines (Grossetti, 2015).

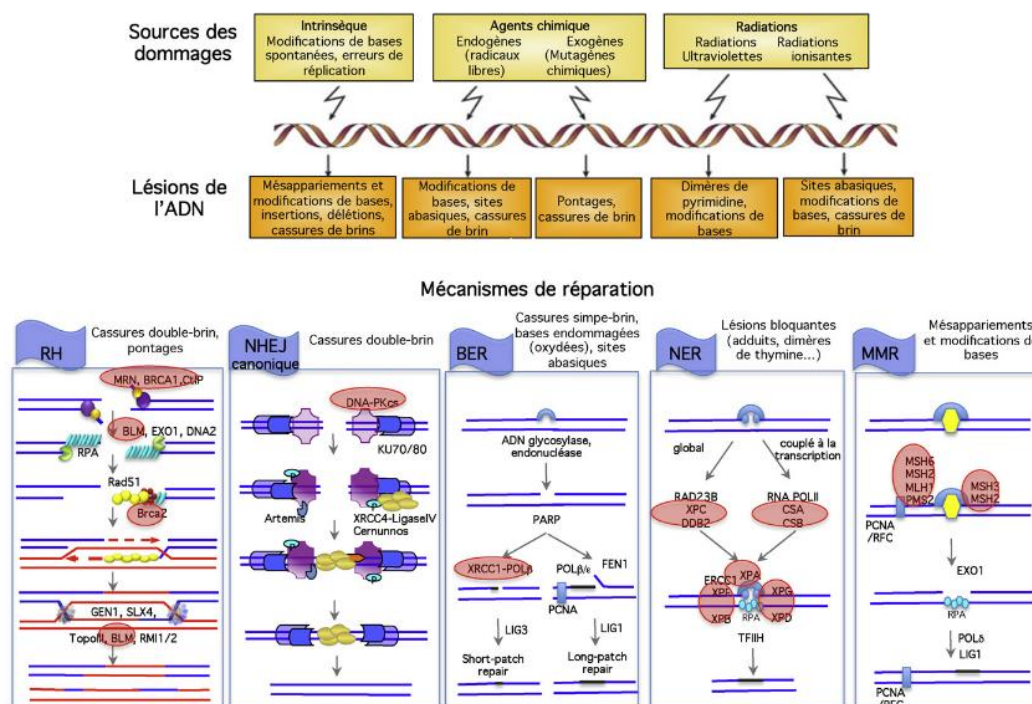


Figure 76 – Différentes lésions et mécanismes chargés de leur réparation¹⁰⁹. Figure extraite de l'article « Réparation des cassures double-brin de l'ADN, un mécanisme peut en cacher un autre : la ligature d'extrémités non homologues alternative. » (Rass, Grabarz, Bertrand, & Lopez, 2012)

¹⁰⁹ La légende entière : « Différentes lésions et mécanismes chargés de leur réparation. Les différentes sources (endogènes ou exogènes) de dommages de l'ADN et les différents types de lésions ainsi que leurs mécanismes de réparation sont schématisés.

La recombinaison homologue est impliquée dans la réparation des cassures double-brin. Avec les protéines de la voie Fanconi (non schématisées ici), la recombinaison homologue joue un rôle crucial dans le redémarrage des fourches de réplication et la réparation des pontages de l'ADN.

La religature d'extrémités non-homologues (NHEJ) permet la réparation des cassures double-brin. L'excision de bases (BER) permet de réparer les dommages de bases tels que les bases oxydées ou les sites abasiques et aussi les cassures simple-chaîne.

La réparation par excision de nucléotide (NER) élimine les lésions bloquantes (tels que les dimères de thymine ou les pontages intra ou inter-brin).

Le système de réparation des mésappariements (MMR) est impliqué dans la réparation des bases mal appariées lors de la réplication (insertion erronée de nucléotide ou glissement (« *slippage* ») de polymérase).

Puisque la terminologie « *DNA repair* » s'est élargie et répandue en même temps que se développaient des sous-groupes spécialisés dans différents mécanismes de réparation de l'ADN, il est intéressant de pousser un peu plus loin l'analyse de la géographie actuelle du domaine. En effet, on pourrait alors voir si le phénomène d'expansion géographique décrit préalablement est associé ou non à la croissance de ces sous-spécialités. Comme nous venons de le montrer, dans plusieurs lieux inactifs au stade d'émergence se trouvent désormais des chercheurs qui font régulièrement référence au terme « *DNA repair* ». Se sont-ils pour autant spécialisés dans l'une des nombreuses branches de la spécialité ? Soit le terme « *DNA repair* » s'est juste banalisé, soit l'activité de recherche interne aux différentes sous-spécialités se diffuse pour de bon. Au-delà, comme la branche de la réparation de l'ADN par excision de nucléotide (NER) est l'une des plus anciennes, il se pourrait que sa diffusion vers de nouveaux lieux ait été rendue difficile (frein à l'entrée) voire pas du tout attractive par rapport à d'autres sous-spécialités.

Pour vérifier ces dernières hypothèses, la moitié des publications de 2006-2008 a été classée à l'aide de mots-clefs par sous-spécialité. Ainsi, l'analyse a montré qu'aucun des 41 pays inactifs avant 1990, a publié plus de 10 « NER » publications en 2006-2008. Au-delà, parmi les 18 agglomérations les plus spécialisées dans le NER (plus de 7 publications dans cette catégorie en 2006-2008), seulement 3 sont entrées tardivement dans l'espace de la réparation de l'ADN (Barcelone, Strasbourg et São Paulo). Mais, puisque cette proportion correspond environ à celle des nouvelles agglomérations à avoir intégré le top 20 des agglomérations publiantes sur la réparation de l'ADN en général (2 entrants tardifs sur 18 pionniers), on peut en déduire qu'il n'y pas plus de frein à l'entrée dans le NER que dans la réparation de l'ADN en général. À Strasbourg, c'est l'évolution interne à la branche NER qui explique cette intégration. Cela vient du fait qu'en 1993, les spécialistes de la transcription à Strasbourg ont participé à la découverte d'un lien entre le mécanisme NER et un facteur de transcription (Chapitre 11).

Dans l'ensemble, contrairement à l'hypothèse initiale, le Tableau 35 montre que davantage d'entrants tardifs ont intégré le NER et le BER (*Base Excision repair*, qui est également un mécanisme de réparation par excision suite à l'apparition d'une lésion sur un seul brin d'ADN) qu'ils n'ont intégré les autres sous-spécialités, et en particulier le groupe DSB spécialisé dans les cassures double brins (*Double Strand Breaks*). Tandis qu'aucun nouveau venu n'a intégré le top 20 des lieux spécialisés dans l'« *Homologous Recombination* » (HR) et le top 20 du groupe DSB en général, on compte 3 agglomérations nouvelles dans le top 20 du groupe BER (Beijing, Novossibirsk, et Gênes) et dans le top

La synthèse trans-lésionnelle impliquant des polymérases mutagènes, connue pour être impliquée dans le développement tumoral, n'est pas représentée ici. Les principales protéines retrouvées déficientes dans les cellules cancéreuses et qui sont associées à la tumorigénèse sont entourées en rouge. » (Rass *et al.*, 2012).

20 du groupe NER (déjà évoquées). En résumé, le niveau de participation des équipes tardivement impliquées dans les mécanismes de réparation par excision (NER et BER) est supérieur au niveau de participation des nouveaux venus dans les autres sous-spécialités (Tableau 35). Comme il est long de se spécialiser et que les mécanismes de réparation par excision étaient prometteurs dans les années 1990, il se peut que les résultats obtenus résultent en fin de compte des événements de ce passé récent.

Les sous-spécialités en réparation de l'ADN et la spécialisation des lieux de publication

		BER*	NER*	MR*	DSB	HR**	Total
Nombre de publications « spécialisées »		444	612	255	1080	475	2404
Agglomérations spécialisées	active en 1965-1975 (112) % de publications	66,69	68,57	74,48	76,63	78,97	72,51
	inactive en 1965-1975 (106) % de publications	33,31	31,43	25,52	23,37	21,03	27,49
<p><i>Agglomérations spécialisées = aires urbaines ayant publié au moins 3 publications spécialisées en 2006-2008</i></p> <p><i>*BER (Base Excision Repair); NER (Nucleotide Excision Repair); MR (Mismatch Repair)</i></p> <p><i>**Homologous Recombination est une sous-spécialité de "DSB" (Double Strand Breaks ou "cassures doubles brins")</i></p>							

Clef: au moins 3 publications spécialisées en 2006–2008 ont été signées depuis 112 agglomérations pionnières (actives en 1965–1975) quand ce niveau de production de publications spécialisées ne concerne que 106 agglomérations entrées plus tardivement dans le domaine de recherche. Ensemble, ces 218 aires urbaines ont publié 2404 publications spécialisées. 72,5% de ces 2404 publications furent produites par les 112 aires pionnières. Cependant, les aires pionnières n'ont participé qu'à 68,5% de la production portant sur le mécanisme de réparation de l'ADN par excision de nucléotide (NER) (612 publications au total). Le tableau montre que les travaux portant sur les cassures doubles brins (DSB) sont plus souvent réalisés dans des aires pionnières que les travaux sur la réparation par excision.

Sources: WoS et Scopus, tous types de documents

Tableau publié en anglais dans l'article « DNA Repair: a changing geography? (1964-2008) publié dans la revue *DNA Repair* en 2013

Tableau 35 – Profil des agglomérations où sont effectuées des recherches spécialisées sur la réparation de l'ADN (« DNA repair »). Source : Scopus et Web of Science

Au bout du compte, les résultats obtenus à l'issue de cette comparaison de la géographie d'un domaine de recherche à son stade d'émergence et de maturité sont en accord avec des résultats connus, par exemple sur l'évolution des réseaux de recherches en biotechnologie. En effet, le travail de Powell sur les réseaux de collaboration entre entreprises de biotechnologie tend, d'après lui, à confirmer un vieux dicton anglais puisque : « tous les premiers entrants ne se transforment pas en gagnants, et certains entrants tardifs arrivent à prendre le dessus. Comme dit le proverbe, le premier oiseau venu attrape peut-être le ver, mais c'est la deuxième souris qui mange le fromage. »¹¹⁰ (W. Powell, White,

¹¹⁰ L'original : « But not all early entrants turn out to be winners, and some latecomers attain prominence. As the saying goes, the early bird may catch the worm, but it is the second mouse that eats the cheese. » (W. Powell, White, Koput, & Owen-Smith, 2005).

Koput, & Owen-Smith, 2005). S'il s'agit d'un phénomène connu en sciences de l'innovation, il n'avait à notre connaissance pas encore été étudié à l'échelle du monde et au niveau d'unités spatiales ; c'est-à-dire que nous en connaissons l'origine sociologique mais pas la traduction géographique.

Dans les résultats ainsi obtenus sur la réparation de l'ADN, les cas les plus flagrants de déclin et d'ascension obéissent à des logiques plus générales ou macro dont certaines, comme la croissance de la science chinoise, avaient précédemment été mises en évidence à partir de la dynamique géographique du système scientifique mondial toutes disciplines confondues. Ceci étant, certaines caractéristiques de la géographie du domaine ne peuvent uniquement s'expliquer par l'état général du système scientifique. Sans tenir compte des logiques scientifiques interne à la réparation de l'ADN, il n'est par exemple pas possible d'expliquer pourquoi une équipe de Strasbourg intègre tardivement la « communauté », ni pourquoi les Pays-Bas et le Texas sont des territoires particulièrement dynamiques dans le domaine depuis ses origines.

Alors que le champ de la réparation de l'ADN s'est complexifié, les évolutions tangibles à l'heure actuelle ont sans doute lieu au niveau des sous-spécialités et dépendent de leurs capacités d'interaction et d'ouverture aux nouvelles opportunités techniques et théoriques en biologie moléculaire, telles que l'étude *in vivo* des fonctions réparatrices ou l'épigénétique (Maisonobe, Giglia-Mari et Eckert, *op. cit.*). Néanmoins, le cadre général et la stabilité que peut offrir une spécialité se veulent structurants pour les développements à venir. C'est en tout cas la conviction qui anime les différents pionniers du domaine lorsqu'ils participent, à travers leur contrôle sur la revue *DNA Repair*, leurs positionnements dans l'espace éditorial global et leurs interventions aux divers événements fédérateurs pour la spécialité, à renforcer sa cohésion et à lui donner une base historique sur laquelle les nouveaux arrivants ont la possibilité de s'appuyer lors de leur intégration.

D'après les récits, l'idée de « communauté » formée par les spécialistes de la « réparation de l'ADN » est née en même temps que le mot-clef s'imposait aux États-Unis et qu'étaient organisés les premières rencontres puis les premiers colloques s'adressant aux chercheurs travaillant exclusivement sur ce thème entre 1965 et 1975. C'est donc sur cette période que nous souhaitons revenir dans la deuxième partie de ce chapitre pour interroger plus en détail le processus d'émergence de la spécialité, et surtout les logiques ayant permis à certains de ses membres de se représenter l'image d'une « communauté internationale de la réparation de l'ADN ». À cette fin, l'analyse s'appuie sur des informations obtenues à l'aide de données formelles, à savoir la localisation des équipes pionnières et les trajectoires bibliographiques de plusieurs pionniers, qui sont croisées avec des informations de nature plus qualitatives recueillies au fil des lectures et des entretiens que nous avons réalisés sur les origines du champ.

2. L'émergence de la spécialité « DNA Repair »

« Comment est-il possible alors de concilier l'accent mis sur l'action locale, contingente et située, avec la continuité des problèmes scientifiques dans le temps, l'espace et les institutions ? » (Keating, Cambrosio, & Mackenzie, 1996)

2.1. Les trajectoires des pionniers de la spécialité

Puisque les contours d'une spécialité sont flous et mouvants, nous avons observé que la perception qu'en ont les chercheurs peut varier d'une équipe à une autre. Ces variations s'expliquent à la fois par l'histoire des équipes et leur localisation, en témoigne certains récits que nous allons retranscrire. Ces récits ont été reconstruits à la suite de six entretiens réalisés en Juin 2011 aux Pays-Bas auprès de chercheurs ayant co-publié des articles faisant référence au mot-clef « *DNA repair* ». Plusieurs équipes de recherche publient aujourd'hui sur la réparation de l'ADN aux Pays-Bas, et ceci depuis plusieurs villes mais surtout Rotterdam et Leyde. La force de frappe du pays dans ce domaine impressionne si bien que Cleaver *et al.* parlent de « *Dutch army* » pour l'évoquer¹¹¹, ce qui ne plaît pas forcément aux intéressés. De son côté, l'un des pionniers du pays, D. Bootsma, qui a longtemps été chef d'équipe à Rotterdam, reconnaît dans une synthèse sur les contributions hollandaises que : « La réparation de l'ADN a été, et est encore, un sujet de recherche important pour un groupe considérable de scientifiques aux Pays-Bas. »¹¹² (Bootsma, 2001).

Tandis que les chercheurs rencontrés à Rotterdam n'ont aucune objection à être considérés comme des spécialistes de la réparation de l'ADN, leurs collaborateurs de Leyde (ville-mère de la revue *Mutation Research*) se disent spécialisés dans la recherche sur les mutations et la dégradation cellulaire. Pour eux, cela n'a pas de sens de se dire « spécialistes de la réparation de l'ADN » puisque les recherches sur les mécanismes de réparation sont intimement liées aux études des lésions sur l'ADN (*DNA damages*). Cette histoire suggère que l'appréhension de la spécialité diffère chez des scientifiques concernés par les mêmes travaux y compris quand ils appartiennent au même pays et au même groupe de

¹¹¹ « Their leadership and impact on the field brings to mind a notable phrase in Sir R. Attenborough's play *Oh What a Lovely War* (film version 1969): "Lord, give me a victory before the American (Dutch?) marines arrive!" » (Cleaver, Karplus, Kashani-Sabet, & Limoli, 2001) .

¹¹² L'original : « DNA repair has been, and still is, an important topic of research of a relatively large group of scientists in The Netherlands. » (Bootsma, 2001).

collaboration ; et cette appréhension n'est évidemment pas stable dans le temps. La littérature existante sur les spécialités suggère que l'appréhension d'un domaine de recherche dépend de l'expérience que l'on en a localement (Gläser & Laudel, 2001), et il se trouve justement que les chercheurs du laboratoire de Leyde et celui de Rotterdam n'ont pas investi la thématique de la réparation avec la même approche et les mêmes intentions.

L'équipe de Rotterdam est, en quelque sorte, le fruit des premières découvertes sur la réparation de l'ADN. Créée en 1972 dans la toute nouvelle université de Rotterdam, elle est dirigée par D. Bootsma qui accepte un poste de choix pour développer, au sein d'une nouvelle équipe, sa thématique de thèse jugée prometteuse en biologie cellulaire. Au contraire, le laboratoire de Leyde est antérieur à celui de Rotterdam et à la mise en évidence de mécanismes de réparation de l'ADN. Dès lors, les recherches qui sont effectuées à Leyde demeurent plus traditionnelles et plus généralistes que celles de Rotterdam ; tandis que la jeune équipe de Rotterdam se montre, dès le départ, plus flexible et ouverte aux nouvelles techniques, comme en témoigne le recrutement d'un « biologiste moléculaire » dès la fin des années 1970.

La situation est encore différente à Amsterdam où c'est essentiellement le mécanisme de réparation par mésappariement (*Mismatch Repair*) qui est étudié, et cela depuis seulement le début des années 1990. Ce mécanisme diffère des autres dans la mesure où il corrige les erreurs produites par l'organisme lui-même, quand les autres réagissent aux dégâts perpétrés par l'exposition du corps aux agressions extérieures : produits chimiques, radiations UV, ionisantes etc. Avant les années 1990, il existait aussi des travaux traitant de l'*Homologous Recombination* (HR) à Amsterdam mais il s'agissait, d'après Hein te Riele (rencontré sur place), d'exploiter les fonctions de ce mécanisme comme outil pour manipuler le matériel génétique, et notamment pour intervenir sur la réplication de l'ADN, mais pas de le prendre pour objet de recherche.

Remarquons que sans notre rencontre avec des protagonistes locaux, il aurait été difficile de saisir les spécificités et les causes de ces différentes perceptions du domaine et d'en faire ici la restitution¹¹³. Ces récits indiquent que si nous souhaitions avoir une appréhension vraiment complète du domaine, il faudrait interroger des protagonistes de tous les sites géographiques concernés, ce qui dépasse les prétentions de ce travail. Ajoutons que des générations de spécialistes des sciences ont attiré l'attention sur la prudence qu'il fallait avoir avec les témoignages des informateurs (par exemple : Mulkay in Lemaine *et al.* 1976). À notre sens, il n'est pas question de démêler le vrai du faux entre la définition de Leyde et celle de Rotterdam, mais de retenir que les travaux entamés dans les années 1960 sur la réparation de l'ADN ont justifié la création d'équipes, qui comme celle de Rotterdam, se sont « spécialisées » volontairement dans la question.

¹¹³ La liste des personnes rencontrées est consultable en Annexe.

C'est ce phénomène d'apparition d'équipes spécialisées dès l'origine du domaine, qui retient notre attention. Notre objectif est de localiser ces équipes spécialisées en différents points du globe et de saisir le rôle que certains pionniers, comme D. Bootsma, ont joué dans l'émergence de ces équipes et leur sentiment d'appartenir à une « communauté ». En nous concentrant sur les trajectoires géographiques des pionniers, nous tâchons de mettre en évidence les réseaux de lieux structurants au stade d'émergence de la spécialité. On considère que ces réseaux de lieux déterminent autant qu'ils sont construits par les trajectoires individuelles. Quatre courants dans la littérature peuvent être mobilisés comme cadres théoriques dans cette perspective : les modèles classiques de diffusion de l'innovation (Rogers, 1962 ; Crane, 1969a), la sociologie des sciences et des connaissances scientifiques (Lemaine *et al.*, 1976 ; Zuckerman, 1988), les modèles évolutionnistes appliqués à l'économie de l'innovation (*evolutionary economics* : Boschma & Frenken, 2006), et enfin les méthodes narratives ou prosopographiques (Lemercier & Picard, 2011 ; Abbott, 2011 ; Bidart & Dupray, 2014). Avant d'en venir à l'élaboration, la restitution et la discussion autour des trajectoires géographiques de quelques pionniers, nous résumons ce que nous avons retenu de cette littérature pour nourrir notre approche.

Si l'on devait synthétiser l'apport de ces quatre courants (diffusion de l'innovation, sociologie des sciences, économie évolutionniste, analyses de trajectoires) pour envisager les logiques d'émergence de la spécialité de la Réparation de l'ADN dans le temps et l'espace, voici ce que nous retiendrions :

1. La communication formelle et informelle entre chercheurs permet la transmission des idées et des savoir-faire nécessaires à l'émergence de spécialités scientifiques. Ce phénomène s'apparente à un processus de diffusion d'innovation à cela près que les produits immatériels qui se diffusent (idées et savoir-faire) se modifient et évoluent en circulant. En constante reconfiguration, l'espace de circulation devient l'espace d'une « communauté » dès lors que la circulation s'organise au sein d'un collectif « globalisé » conscient de partager des intérêts communs.
2. Les pratiques scientifiques sont situées, en particulier au niveau des laboratoires. C'est dans ce cadre que les biologistes moléculaires sont amenés à agir et travailler quotidiennement. La variabilité des profils de spécialisation d'un laboratoire à un autre s'explique par la diversité des collectifs « localisés » (équipes) qui peuvent s'y former et s'y déformer. Au sein de leurs équipes et dans le laboratoire, les scientifiques ont des statuts, des stratégies et des profils variés qui orientent leurs actions et leur point de vue (Knorr-Cetina, 1982).

3. S'appuyant sur le principe de la « dépendance du sentier » (*path dependency*), la théorie évolutionniste suggère que l'absorption d'une nouveauté en un lieu donné dépend de ce qui était présent dans ce lieu auparavant et, en ce sens, ne dépend pas essentiellement de la volonté individuelle. Suivant cette logique, le profil de spécialisation d'un laboratoire n'est pas déconnecté des spécialisations présentes auparavant ou simultanément dans ce laboratoire. C'est ce que nous avons observé dans le cas du laboratoire de Leyde, voire même dans celui d'Amsterdam.
4. En revanche, « la dépendance du sentier » ne nous permet pas d'expliquer la spécialisation du laboratoire de Rotterdam puisque ce dernier a été créé de toutes pièces au sein d'une nouvelle université. L'approche narrative, de son côté, nous invite à considérer les trajectoires individuelles et les facteurs institutionnels, comme l'ont fait par exemple Michel Grossetti et Pierre Mounier-Kuhn pour expliquer la carte de l'émergence de l'informatique en France (Grossetti & Mounier-Kuhn, 1995) ou encore Bruno J. Strasser pour comprendre l'apparition de la biologie moléculaire à Genève (Strasser, 2002). En effet, le développement des recherches sur la réparation de l'ADN à Rotterdam ne se comprend que si l'on considère la trajectoire de D. Bootsma et les facteurs institutionnels ayant permis sa mobilité. Conduit à quitter le centre de recherche où il avait fait sa thèse, à savoir le centre national de biologie médicale du TNO (Organisation Nationale de Recherche Appliquée) de Rijswijk pour fonder une équipe dans l'Université de Rotterdam, on peut imaginer que le succès de cette opération n'aurait pas été possible si l'équipe de radiobiologie de Rijswijk n'avait pas été victime de coupes budgétaires et encouragée à se consacrer à des recherches plus appliquées.

Les deux derniers points sont essentiels pour notre question, tandis que les deux premiers sont importants à garder en mémoire pour l'interprétation des données que nous analysons. En effet, les troisième et quatrième propositions sont au fondement de l'étude de trajectoires qui va suivre mêlant analyses bibliométrique, prosopographique et narrative. Le travail de reconstitution des trajectoires s'articule en deux temps. Premièrement, il s'agit de revenir sur le corpus des publications « *DNA repair* » parues au stade d'émergence de la spécialité, c'est-à-dire le corpus extrait de *Scopus*. En s'appuyant sur ce corpus, on identifie un ensemble de villes, d'institutions et de chercheurs dont le niveau d'activité sur le thème de la réparation de l'ADN est jugé significatif. L'idée sous-jacente est qu'à ce stade où le mot-clef vient juste de faire son apparition, l'utiliser à plusieurs reprises est à la fois la preuve d'avoir été en contact avec les tous premiers articles publiés sur ce thème, et de vouloir s'en rapprocher par le biais de ses propres recherches. Après

avoir ainsi repéré un groupe de chercheurs pionniers, on s'intéresse à leur trajectoire bibliométrique, c'est-à-dire aux laboratoires à partir desquels ils ont publié des articles entre 1960 et 1975. On verra alors si ces trajectoires parviennent comme dans le cas de D. Bootsma à expliquer l'expansion géographique de la spécialité au stade d'émergence, et dans quelle mesure elles ne le peuvent pas ou imparfaitement.

À partir du corpus, on repère 42 agglomérations urbaines d'où ont été signés au moins 10 articles avec le mot-clef « DNA Repair » entre 1965 et 1975 (Tableau 36) :

Les agglomérations urbaines d'où proviennent au moins 10 publications "DNA Repair" en 1965-1975

Baie de San Francisco Californie	Philadelphie Pennsylvanie	Rome Italie
Moscou Russie	Vancouver Bc Canada	Houston Texas
New-York	Manchester Royaume-Uni	Oslo Norvège
Knoxville Tennessee	Vienne Autriche	Ithaca Etat de New-York
Londres Royaume-Uni	Dallas Texas	Birmingham Alabama
Boston Massachusetts	New-Haven Connecticut	Mishima Shizuoka Japon
Tokyo Japon	Pushchino Russie	Baltimore Maryland
Bethesda Maryland	Toronto On Canada	Rochester Etat de New-York
Leyde-Rotterdam Pays-Bas	Durham Caroline du Nord	Edimbourg Royaume-Uni
Los-Angeles Californie	Berlin Allemagne	Heidelberg Allemagne
St-Petersbourg Russie	Gainesville Floride	Fukuoka Fukuoka Japon
Brighton Royaume-Uni	Bratislava Slovaquie	Bruxelles Belgique
Kyoto Japon	Madison Wisconsin	
Paris France	Francfort Allemagne	
Chicago Illinois	Stockholm Suède	

Tableau 36 – Liste des agglomérations urbaines spécialisées dans la réparation de l'ADN entre 1965 et 1975. Source : Scopus

Ensuite, après l'harmonisation intégrale du champ « adresse », on obtient la liste des institutions scientifiques d'où furent signés au moins 10 articles « DNA repair » entre 1965 et 1975 (comptage entier non fractionné). Elles sont 48 (Tableau 37). Dans cette liste, on remarque qu'il n'y a pas d'institutions allemandes, scandinaves et italiennes alors que Berlin, Francfort, Heidelberg, Stockholm, Oslo et Rome figurent parmi les agglomérations « spécialisées » (Tableau 36). Cela indique que les publications provenant de ces agglomérations sont le fait d'institutions, d'équipes et d'auteurs « non spécialisés » d'après le critère de spécialisation que nous avons retenu. Il faudrait en effet que les institutions et les auteurs à avoir publié sur la réparation de l'ADN depuis ces agglomérations aient eu une activité de production répétée sur la question pour pouvoir être repérés comme « spécialistes », ce qui n'est pas le cas.

Laboratoires d'où proviennent au moins 10 publications "DNA Repair" en 1965-1975

Stanford Univ., US	Harvard Univ., US	Johns Hopkins Univ., US
Oak Ridge Nat. Lab. & Tennessee Univ.	Univ. Texas, Dallas, Richardson, US	Univ. Chicago, US
Univ. California, San Francisco	Univ. Tokyo, Japon	Cornell Univ., Ithaca, US
Univ. Sussex, Brighton, UK	University of California, Berkeley	Brandeis Univ., US
NCI, NIH, Bethesda, MD, US	Temple Univ., US	Nat. Inst. Genet., Mishima, Japon
Acad. Sci. USSR, Leningrad, Russie	Argonne Nat. Lab., IL, US	Univ. Alabama, US
Acad. Sci. USSR, Moscou, Russie	Forsch. Zent., Seibersdorf, Autriche	Univ. Rochester, US
ICR, London, UK	Univ. Florida, Gainesville, US	Inst. Environm. Hlth Sci., Research Triangle Park
USSR Acad. Med. Sci., Moscou	Brookhaven Nat. Lab., US	Michigan State Univ., East Lansing, US
Univ. British Columbia, Vancouver	Nat. Canc. Cent. Res. Inst., Japon	Atom. Energy Canada Ltd, Chalk River
Yale Univ., US	Univ. Wisconsin, Madison, US	Free Univ Bruxelles
UCLA, Los Angeles, US	Cancer Res. Inst., Slovak Acad. Sci.	Inst. Atom. Energy, Moscow, Russie
Leiden State Univ., NDL	Fond. Curie inst radium, France	Kyushu Univ., Japon
Osaka Univ., Japon	Rutgers Univ., New Brunswick, US	USSR Acad. Med. Sci., Obninsk, Russie
Christie Hosp., Holt Radium Inst., UK	TNO Rijswijk, Pays-Bas	York Univ., Toronto, Canada
Acad. Sci., Pushchino, Russie	Colorado State Univ., Fort Collins, US	

Tableau 37 – Liste des institutions spécialisées dans la réparation de l'ADN entre 1965 et 1975. Source : Scopus

Enfin, dans l'intention de saisir les circulations entre les premiers laboratoires spécialisés, on augmente encore la focale pour se concentrer sur la population des quelques auteurs à avoir publié au moins 10 articles « *DNA repair* » entre 1965 et 1975 (comptage entier non fractionné). En faisant le choix d'utiliser régulièrement le mot-clef dès son apparition, ces auteurs ont contribué à faire exister pour la première fois la dite « spécialité » en différents lieux. Après l'harmonisation du champ « auteur » dans le corpus, 35 auteurs que l'on qualifie de « pionniers » sont ainsi identifiés (Tableau 38).

Auteurs ayant signés au moins 10 publications "DNA Repair" en 1965-1975

Smith, K.C.	Skavronskaya, A.G.	Elkind, M.M.
Cleaver, J.E.	Witkin, E.M.	Harm, W.
Bridges, B.A.	Eberl, R.	Sedgwick, S.G.
Stich, H.F.	Kondo, S.	Byfield, J.E.
Setlow, R.B.	Regan, J.D.	Fox, B.W.
Painter, R.B.	Roberts, J.J.	Friedberg, E.C.
Altmann, H.	Setlow, J.K.	Klein, W.
Hanawalt, P.C.	Trosko, J.E.	Robbins, J.H.
Fox, M.	Wu, R.	Sedliakova, M.
Lieberman, M.W.	Bootsma, D.	Youngs, D.A.
Zasukhina, G.D.	Cerutti, P.A.	Paterson, M.C.
Gaziev, A.I.	Dubin, N.P.	

Tableau 38 – Liste des pionniers de la spécialité ayant signé au moins 10 publications « *DNA repair* » entre 1965 et 1975. Source : Scopus

Passée cette étape, l'enjeu consiste à reconstituer la mobilité institutionnelle de ces auteurs à partir des informations bibliométriques. À chacun des 35 auteurs, sont associées une ou plusieurs affiliations. Il s'agit de renseigner pour chaque auteur, l'ensemble des affiliations qu'il a occupées depuis 1960, l'ordre d'occupation et sa durée. Ces informations sont, avec les imprécisions déjà évoquées, accessibles dans la base *Scopus* puisqu'à chaque article publié, les auteurs signalent une affiliation qui est indexée avec l'article dans la base. Pour chaque auteur, nous avons donc systématiquement recensé l'ordre d'apparition des affiliations qu'il a occupé à partir de 1960 et leur durée.

Comme on pouvait s'y attendre, les 35 pionniers ne peuvent à eux seuls expliquer l'ensemble du phénomène d'émergence. Ensemble, ils ont écrit 397 publications « *DNA repair* » entre 1965 et 1975, soit un peu plus de 20 % de l'ensemble des publications parues à cette période. Entre 1960 et 1975, ils sont passés par 33 institutions « spécialisées » et 25 agglomérations urbaines d'importance pour la spécialité. Enfin, les lieux les plus fréquentés par ces auteurs se trouvent pour les 2/3 (14 agglomérations sur 25) aux États-Unis. Avant toute interprétation, il faut se rappeler que l'information n'est pas de qualité égale dans la base. En parcourant la base pour reconstituer les trajectoires institutionnelles des « pionniers », nous avons observé que les notices des articles publiées dans des revues américaines étaient mieux et plus systématiquement renseignées que les notices d'articles des autres revues, et notamment des revues soviétiques.

Afin de synthétiser les informations recueillies sur les carrières des auteurs, on a opté pour une représentation graphique sous la forme d'un graphe bipartite (*two-mode*) auteurs-institutions. Pour garder une trace de l'ordre d'occupation des institutions, la couleur des liens dépend de l'ancienneté de la relation : les liens noirs sont ceux qui lient l'auteur à la plus ancienne des institutions dans laquelle il a travaillé ; et les liens gris, à la dernière institution fréquentée avant 1975. La représentation graphique permet ainsi de mettre en évidence d'intéressants faits longitudinaux (Figure 77).

En premier lieu, on constate que le réseau est doté d'une grande composante principale à laquelle appartiennent 14 des 35 auteurs dont nous avons suivi la trajectoire. On peut d'ores et déjà présumer qu'il s'agit là d'un groupe auteurs-institutions central pour l'émergence de la spécialité. Cet ensemble regroupe des institutions en majorité américaines, et des institutions anglaises, néerlandaises et canadiennes y sont aussi rattachées (l'Université du Sussex et le *Medical Research Center* à Brighton ; l'Université de Leyde ; le Centre de recherche en énergie atomique de Chalk-River au Canada et celui de Harwell en Grande-Bretagne).

En dehors de cet espace d'interaction dense, on compte plusieurs composantes secondaires. Parmi elles : 9 chercheurs sont isolés et ne partagent d'institution avec aucun autre chercheur ; 3 spécialistes sont à Moscou mais ne changent pas d'adresse entre 1960

et 1975 ; un groupe de co-auteurs est en Autriche ; 2 chercheurs (le couple Fox) sont à Manchester tout au long de la période et enfin, bien que proche de l'Université de Leyde attachée à la composante principale (par le biais de la mobilité de MC Paterson), la trajectoire de D. Bootsma en est indépendante¹¹⁴.

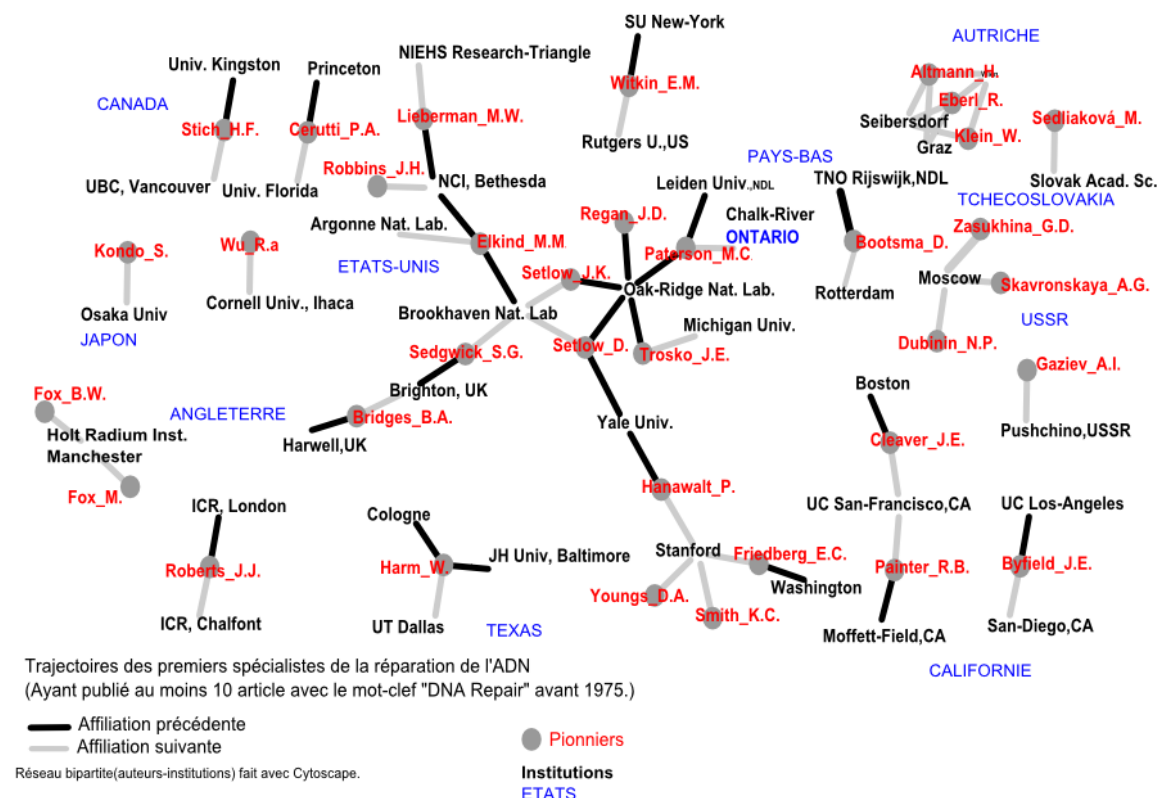


Figure 77 – Graphe bipartite auteurs-institutions synthétisant les parcours institutionnels suivis par 35 auteurs pionniers dans l'utilisation du mot-clef « DNA repair » entre 1960 et 1975. Extrait de l'article « Émergence d'une spécialité scientifique dans l'espace géographique - La réparation de l'ADN » (Maisonobe, 2015)

Dans l'ensemble, on remarque que les chercheurs du bloc de l'Est (Russie et Tchécoslovaquie) ne sont pas concernés par la mobilité institutionnelle. En effet, de ce côté-là du monde, on ne trouve de traces de mobilité ni avec l'international ni à l'intérieur du bloc de l'Est.

Détail important pour nous : la France et la Belgique sont absentes de la représentation alors que l'Institut du Radium à Saclay et l'Université Libre de Bruxelles figurent parmi les institutions « spécialisées » (Tableau 37). Peut-être est-ce à cause de la barrière linguistique, mais cela signifie qu'il n'y a pas dans ces pays de chercheurs ayant signé au

¹¹⁴ Pour une meilleure lisibilité et dans la mesure des possibles, nous avons choisi de positionner les composantes en fonction de leur plus ou moins grande proximité géographique.

moins 10 publications « *DNA repair* » dans des revues indexées par *Scopus* entre 1965 et 1975. Par conséquent, les pionniers francophones du domaine, sur lesquels on reviendra, ne figurent pas sur cette représentation.

Pour pouvoir discuter ou commenter la représentation plus avant, les informations biographiques recueillis au sujet des chercheurs de la réparation de l'ADN, les synthèses historiques écrites par ces derniers, et les entretiens que nous avons menés aux Pays-Bas gagnent à être mobilisés. La représentation fait en quelque sorte office de support pour la narration.

2.2. *La rencontre entre le mythe et les premières traces d'activité*

La plupart des écrits qui reviennent sur l'histoire de la réparation de l'ADN mettent en avant D. Setlow (Oak Ridge) et P. Hanawalt (Stanford) qui ont prouvé l'existence de la réparation de l'ADN par excision de nucléotide (NER) en 1964. Il s'agit d'un mécanisme permettant de réparer les lésions simple brin sur l'ADN provoquées par des agents extérieurs comme les rayons UV. Les témoignages de D. Setlow et P. Hanawalt sur cette période racontent leurs expériences et la teneur de leurs échanges avant qu'ils publient indépendamment leurs résultats en 1964, en officialisant par la même occasion le recours au mot-clef « *DNA repair* » (Philip Hanawalt, 2010 ; Setlow, 2005). Après 1964, P. Hanawalt et D. Setlow jouent un rôle important en formant des scientifiques et en organisant des rencontres. Rien d'étonnant alors à ce que la composante principale du réseau de trajectoires s'articule autour de ces deux personnages (Figure 77). Comme on peut l'observer sur la représentation, ils sont tous deux issus du département de biophysique de l'Université de Yale, où Setlow raconte avoir été le « *mentor* » d'Hanawalt. Les lieux où ils ont passé la plus grande partie de leur carrière, Stanford et Oak Ridge sont les lieux les plus connectés de la figure, c'est-à-dire, ceux par lesquels sont passés le plus de pionniers de notre échantillon.

Les preuves du NER en 1964 génèrent un enthousiasme parmi les photobiologistes et radiobiologistes, qui s'intensifie en 1968 quand le post-doctorant J. Cleaver découvre l'existence d'un lien entre une maladie orpheline (*Xeroderma Pigmentosum* alias « la maladie de la lune ») et un dysfonctionnement du NER provoquant des cancers de la peau. Cette découverte a lieu dans le nouveau laboratoire de radiobiologie de UCSF dirigé par R. Painter. Pour en assurer la direction, ce dernier a tout juste quitté le centre de recherche de la NASA situé à Moffett-Field aux alentours de San Francisco (*Ames Research Center*). D'après Hanawalt, en établissant un lien entre les mécanismes de réparation de l'ADN et le cancer, la découverte de Cleaver en 1968 place le nouveau champ de recherche en « orbite » et débloque des financements. Sur la Figure 77, la composante

« Cleaver - Université de San Francisco - Painter » est indépendante de la composante principale. Néanmoins, nous savons, grâce aux récits d'Hanawalt, qu'en plus de partager une proximité géographique, Hanawalt et Painter se connaissaient (Hanawalt, 2003).

Les récits des Américains s'accordent à dire que l'une des premières grandes rencontres internationales sur la réparation de l'ADN fut organisée en 1974 par Hanawalt et Setlow à Squaw Valley en Californie. Elle aurait attiré 200 participants. Ayant occupé son premier poste à Stanford, Friedberg se souvient que Stanford est alors le siège d'« une communauté de la réparation de l'ADN en miniature » avec entre autres : K.C. Smith au département de radiologie, Hanawalt au département de biologie ou encore A. Kornberg au département de biochimie (Friedberg, 2005). D'une autre nature et de l'autre côté du pays, le lieu de travail de Setlow est tout aussi structurant. Il s'agit de la division de biologie du laboratoire national d'Oak Ridge (Tennessee), dans laquelle exercent Richard et Jane Setlow. Un très grand nombre de jeunes chercheurs viennent s'y former. D'après Setlow : « Grâce aux efforts de son directeur Alexander Hollaender, c'était l'un des plus grands centres de biologie du monde. Il avait du personnel et de l'expertise en biophysique, biochimie, génétique, immunologie, pathologie, carcinogénèse etc. et des équipements de qualité dans tous les domaines. »¹¹⁵ (Setlow, 2005). Les travaux de l'historienne Angela Creager témoignent aussi très bien du rôle structurant d'Oak Ridge, ex laboratoire du projet Manhattan, pour le tournant médical des études sur les radiations (voir par exemple : Creager et Santesmases, 2006).

Puisqu'on y trouve les premiers pionniers à avoir lancé le mot-clef « *DNA repair* », il n'est pas étonnant qu'Oak Ridge et Stanford soient les lieux les plus connectés du réseau, avec en plus le laboratoire de Brookhaven situé sur Long Island où s'est rendu le couple Setlow suite aux coupes budgétaires intervenues à Oak Ridge. En plus d'être centraux, la figure montre qu'Oak Ridge et Brookhaven ont accueilli des chercheurs venus de l'étranger : M.C. Paterson et S.G. Sedgwick. Un seul autre laboratoire est dans ce cas qui se trouve à l'Université Johns Hopkins de Baltimore. C'est là-bas que Walter Harm a atterri, on y reviendra. Étant les seules trajectoires transnationales visible sur le réseau, les trajectoires de Paterson, Sedgwick et Harm méritent qu'on s'y arrête.

1. Malcolm C. Paterson (flèche verte sur la Figure 78) est un chercheur d'origine canadienne qui fut l'un des premiers doctorants de Richard Setlow à Oak Ridge. Avant d'obtenir un poste de chercheur au laboratoire de recherche nucléaire de

¹¹⁵ L'original : « As a result of the efforts of the Division Director, Alexander Hollaender, it was one of the largest biology research centers in the world. It had scientific personnel and expertise in biophysics, biochemistry, genetics, immunology, pathology, carcinogenesis, etc. and state-of-the-art equipment in all fields, that included a high-energy monochromator. » (Setlow, 2005).

Chalk River (Ontario, Canada), Paterson va à la rencontre des pionniers néerlandais à l'occasion d'un séjour post-doctoral. En Hollande, il est affilié au laboratoire de Leyde mais il co-écrit avec P.H. Lohman, alors à la tête de l'équipe de Rijswijk¹¹⁶. Plus tard, probablement suite à un changement de conjoncture à Chalk River, Paterson part pour l'Alberta où il intègre un institut de recherche sur le cancer à Edmonton au sein duquel il monte un programme d'oncologie moléculaire. Il sera encore très mobile dans la suite de sa carrière puisque ses affiliations permettent de le localiser en Arabie Saoudite (Riyadh) et à Singapour avant qu'il ne prenne sa retraite en Colombie Britannique¹¹⁷.

2. Steven G. Sedgwick (flèche bleue sur la Figure 78) est un chercheur anglais ayant fait sa thèse dans l'unité de recherche sur les mutation cellulaire du *Medical Research Council* (MRC) associé à l'Université du Sussex (Brighton). Le directeur et créateur de cette unité était Bryn Bridges qui avait démarré sa carrière comme radiobiologiste dans l'unité du MRC d'Harwell. Après sa thèse, Sedgwick obtient des bourses non seulement pour venir à Brookhaven mais aussi pour aller à Oak Ridge et en France (dans l'équipe de Raymond Devoret au laboratoire d'enzymologie du CNRS à Gif sur Yvette). Après tout ces déplacements, en 1978, Sedgwick est recruté à la division génétique de l'institut national de recherche médicale à Mill Hill (NIMR, grand Londres). À ce propos, il se souvient : « J'ai organisé la première rencontre du réseau de la réparation de l'ADN au NIMR en 1979. À l'époque, c'était une rencontre isolée, mais j'ai ensuite développé un programme de rencontres, dont je me suis chargé jusqu'au début des années 1990. L'idée de fond était d'encourager les jeunes scientifiques à se lancer dans le domaine et cela continue encore de fonctionner aujourd'hui. »¹¹⁸. Comme aucun article n'a été signé par Sedgwick à Oak Ridge et que les articles qu'il a co-écrits à Gif-sur-Yvette sont seulement parus en 1978, il n'était pas possible de saisir ces autres déplacements dans le réseau des trajectoires de 1960-1975 ; ce qui permet au passage de souligner les limites de notre approche, on y reviendra.

¹¹⁶ En fait, Lohman et Paterson ont fait leur post-doctorat en même temps aux Etats-Unis, l'un au Texas, le second dans le Tennessee, sur un sujet proche (Source : entretien avec P. H. Lohman en 2011).

¹¹⁷Source : URL: <http://www.aihealthsolutions.ca/researchnews/2008/fall/followingup/>

¹¹⁸ L'original : « I organised the first DNA Repair network meeting at NIMR in 1979. At the time, it was a one-off meeting but then it developed into a programme of meetings which I ran until the early 1990s. The original emphasis was to encourage more junior scientists and this theme has continued to the present day. ». Source : URL: <http://www.nimr.mrc.ac.uk/news/steve-sedgwick-awarded-inaugural-genome-stability-network-medal/>.

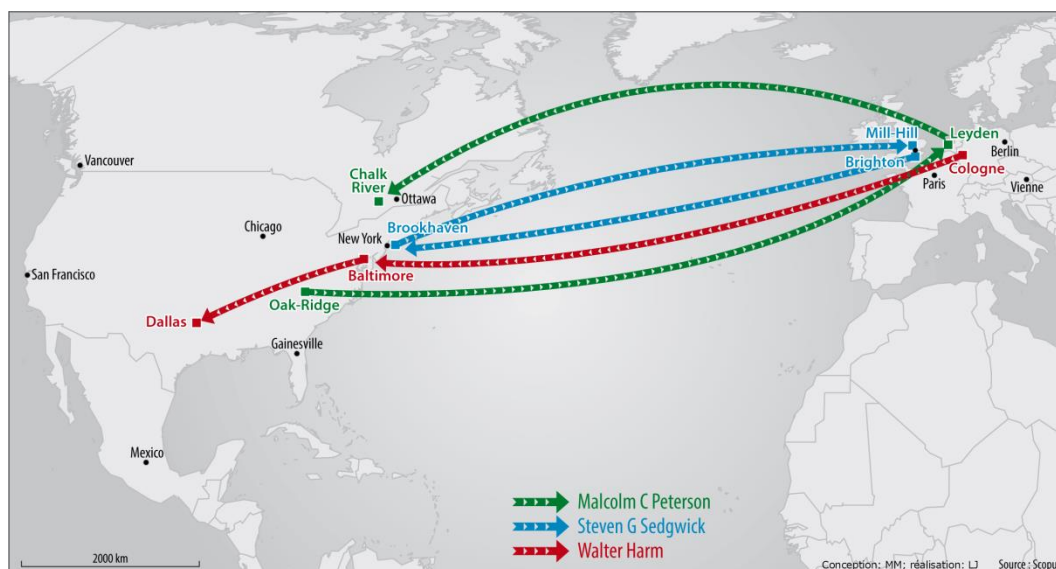


Figure 78 – Les trajectoires internationales des pionniers de la réparation de l'ADN (1960-1975). Source : Scopus et Web of Science

Enfin, Walter Harm (flèche rouge sur la Figure 78) est un chercheur allemand qui a travaillé avec sa femme Helga et Stan Rupert à l'Université Johns Hopkins de Baltimore avant d'être recruté au milieu des années 1960 par Carsten Bresch (un autre chercheur allemand) au *Southwest Institute for Advanced Studies* (SCAS) du Texas. En fait, Bresch et Harm avaient été collègues à l'Institut de Génétique de Cologne, créé par le fameux Max Delbrück. Dans son livre, Friedberg raconte cette histoire parce que le groupe allemand, Delbrück inclus, a joué un rôle important dans la découverte de la photoréactivation. En plus, comme Friedberg travaille à Dallas depuis le début des années 1990 (il a fait le début de sa carrière à Stanford au côté d'Hanawalt), il a bien connu les protagonistes de cette histoire. Ainsi, Stan Rupert aurait rencontré Walter Harm pour la première fois à Cologne. Comme il était en année sabbatique à Copenhague, Rupert y avait été convié par Delbrück pour donner un séminaire. Quand après deux ans à la tête de l'institut de Cologne, Delbrück finit par regagner les États-Unis, « il laissa un très grand vide, et l'institut se trouva aussitôt en crise. »¹¹⁹. En 1965, Bresch quitte donc l'institut pour prendre la tête de la division de Biologie du SCAS à Dallas. Fondé par Texas Instrument en 1948, le SCAS a ensuite été repris par l'État du Texas en 1969 pour devenir l'Université du Texas à Dallas. Walter Harm y exerce pour le reste de sa carrière dans le groupe des photobiologistes ou biophysiciens des UV. D'après Friedberg, « Le Southwest Institute/UT Dallas était peut-être le début de ce qui a été plus récemment qualifié de

¹¹⁹ L'original : « he left a huge leadership vacuum, and the institute fell into crisis. » (Friedberg, 1997, p. 50-56).

"*Texas Mafia*" dans le champ de la réparation de l'ADN, avec d'importants groupes établis à la fois à Dallas, San Antonio, Houston, et Galveston. »¹²⁰ (Friedberg, 1997, p. 50-56).

Ensemble, ces trois scientifiques, pas les plus fameux de la communauté, sont passés par quelques uns des lieux les plus centraux de la spécialité au moment de sa genèse. En traversant les frontières, et surtout l'Atlantique, ils ont contribué à établir ou entretenir un réseau de lieux à l'échelle mondiale. Les détails de leurs trajectoires suggèrent l'existence de liens entre américains, anglais, canadiens, néerlandais, allemands, et français dès les origines du champ. Cependant, tous les autres changements d'affiliations détectés dans notre échantillon ont eu lieu à l'intérieur des cadres nationaux et pratiquement l'intégralité d'entre eux se sont produits aux États-Unis. Outre les biais déjà évoqués, cela vient du fait que les bifurcations repérables à l'aide des données bibliographiques ne mettent en évidence que les mobilités professionnelles durables, et beaucoup plus rarement les courts séjours de travail (qui ne donnent pas forcément lieu à une publication signée depuis le laboratoire d'accueil).

Le fait que les changements d'affiliation en cours de carrières soient fréquents aux États-Unis est cohérent avec les résultats obtenus par Warren Hagstrom lors de ses recherches sur la « communauté scientifique » (Hagstrom, 1965). Selon lui, la dispersion géographique des spécialistes d'un domaine lors de son émergence est un bon moyen de limiter la concurrence au niveau local. De plus, cela permet à la jeune spécialité de se répandre et donc d'être plus rapidement reconnue : le cas de Setlow et Hanawalt ayant quitté Yale pour des localités diamétralement opposées sur le sol américain est sans doute assez révélateur de cette stratégie. Hagstrom reconnaît que ce mode de diffusion est assez courant aux États-Unis, mais qu'il en est sans doute autrement en URSS : « La concentration a plus de chance d'arriver lorsque les équipements de recherche ne peuvent se diffuser. Aussi, elle arrive sans doute plus souvent dans les nations où l'enseignement supérieur est très centralisé et où la recherche est planifiée, comme en Angleterre, en France et en Union Soviétique. »¹²¹ (*ibid.*, p. 167). Les données recueillies sur la réparation de l'ADN tendent à confirmer cette hypothèse, et suggèrent que le rôle de la mobilité pour le développement d'une spécialité varie d'un pays à l'autre. Dans le cas des Pays-Bas, D. Bootsma nous accorde qu'il y a eu une part de planification dans l'organisation du champ : les spécialistes néerlandais ont décidé de diviser leurs forces, et une fois répartis en différents points du territoire, d'éviter la concurrence directe. Cela

¹²⁰ L'original : « The Southwest Institute/UT Dallas group was perhaps the beginning of what has more recently been referred to as the "Texas Mafia" in DNA Repair, with prominent research groups now established in Dallas, San Antonio, Smithville, Houston and Galveston. ». (Friedberg, 1997, p. 50-56)

¹²¹ L'original : « Concentration seems most likely to occur when research facilities cannot be distributed. In addition, it may occur to a greater extent in nations in which university education is highly centralized or research is planned, such as England, France and the Soviet Union. » (Hagstrom, 1965).

supposait de ne pas travailler sur les mêmes maladies, ou au moins sur les mêmes variantes des maladies ; tout en travaillant sur des problèmes suffisamment proches pour être complémentaires et pouvoir collaborer (voir les extraits d'entretiens en Annexe).

En dehors du facteur stratégique, les déplacements observés sur la Figure 77 vont dans le sens des observations de la première partie du chapitre qui montraient le déclin progressif des laboratoires de recherche nationaux sur les radiations comme lieux d'exercice de la spécialité au profit de son développement dans le milieu académique et médical. En effet, le laboratoire d'Oak Ridge de même que le TNO à Rijswijk (Pays-Bas), le Centre de Chalk-River au Canada, et celui d'Harwell dans le Berkshire (Grande-Bretagne), visibles sur la figure, étaient des centres nationaux largement soutenus par l'armée après-guerre pour étudier et développer l'utilisation pacifique de l'énergie atomique à des fins médicales, industrielles et de sécurité intérieure. À partir des années 1970, il n'est plus question d'y financer une recherche fondamentale, tandis que la réparation de l'ADN devient un sujet un peu plus légitime aux yeux des biologistes et généticiens « classiques ». Cette évolution permet d'expliquer les déplacements de D. Bootsma, B. Bridges et J. Trosko qui se traduisent par la création d'équipes spécialisées dans la réparation de l'ADN dans les universités de Brighton, Rotterdam et East-Lansing. Aussi, suite à l'établissement d'un rapport entre le NER et le cancer de la peau en 1968, des équipes se développent dans les centres de recherche sur le cancer comme le *National Cancer Institute* (NCI) de Bethesda (visible sur la Figure 77). La thématique continuera d'y occuper une place importante sachant que petit à petit des liens sont mis à jour avec d'autres cancers ainsi qu'avec le vieillissement.

Ces quelques histoires confirment la pertinence de la représentation graphique obtenue à partir de l'information bibliométrique. En suivant la trajectoire des auteurs sur le graphe, nous retrouvons les connexions, évoquées dans les récits des origines, entre les principaux foyers de la spécialité et pouvons retracer la genèse de plusieurs équipes spécialisées apparues dans les années 1970. Cependant, quand on se fie à la représentation et aux récits des biologistes américains, il semble que l'essentiel de l'action se déroule sur le sol américain. Or, à l'issue de notre rencontre avec K. Jaspers et D. Bootsma à l'Université de Rotterdam en Juin 2011, la situation est moins évidente. Pour Jaspers, le domaine « *Excision Repair* » est un domaine ayant la particularité d'être dominé par les chercheurs européens et notamment néerlandais, et Bootsma et ses collègues considèrent que c'est au Pays-Bas que s'est tenue la première grande rencontre internationale sur la réparation de l'ADN dans les années 1970 (Hoeijmakers, 2001). De plus, les informations bibliométriques suggèrent qu'il se passe également des choses en Russie, en Autriche et au Japon. Or, la contribution des auteurs de ces pays à l'émergence de la spécialité est plus rarement évoquée dans les récits qui participent à fonder la mythologie du domaine. En

suivant les pistes indiquées par les informations bibliométriques, nous avons pu néanmoins en retrouver quelques traces.

S'ils sont actifs sur la question, la Figure 77 suggère que les chercheurs soviétiques sont isolés. Pour autant, d'après les observations des Setlows sur les publications soviétiques, certains chercheurs de l'Est semblent connaître et s'intéresser à la littérature occidentale : « Les généticiens soviétiques sont apparemment divisés en deux camps : ceux qui croient au "Darwinisme créatif" (comme les Lysenkistes) et les individus qui, comme Dubinin, sont influencés par le travail de l'Ouest »¹²² (Setlow & Setlow, 1961). Parmi ces derniers, il y en a même qui arrivent à publier dans les revues occidentales, comme l'illustre l'article paru dans *Mutation Research* en 1971 co-signé par A.G Skavronskaya et Yu S. Kondratiev (Institut Gamelaya d'épidémiologie et de microbiologie de l'Académie des sciences médicales de Moscou). C'est le premier article russe publié avec le mot-clef « *DNA repair* ». Le fait donc qu'ils utilisent le mot-clef et qu'ils publient dans la première revue à se spécialiser dans la question est révélateur de leur volonté de participer à la « communauté ». Dans le même temps, ils ne sont pas complètement ignorés par celle-ci. En effet, si l'article de A.G. Skavronskaya dans *Mutation Research* n'a pas été cité (d'après le WoS), son article paru la même année dans la revue *Molecular and General Genetics* est cité 21 fois entre 1971 et 1975. Parmi les 21 citations, 15 sont localisables : 7 viennent d'URSS, 5 des États-Unis (dont 1 de P. Hanawalt), 1 des Pays-Bas, d'Angleterre et d'Australie. Auparavant, N. Dubinin avait aussi publié dans *Mutation Research*, en 1969, un article qui avait reçu 24 citations avant 1975 ; dont 8 localisables : 3 des États-Unis, 3 d'Italie, 1 d'URSS et 1 du Japon. Malgré l'incomplétude de ces informations, cela suggère que le travail des chercheurs soviétiques est connu et reconnu par les pionniers de la réparation de l'ADN.

D'après Friedberg, les Setlow se sont même rendus en URSS en 1972 à l'occasion du Congrès International de Biophysique. Ils rapportent y avoir fait la connaissance d'Evelyn Waldstein qu'ils auraient convié à venir visiter le laboratoire d'Oak Ridge « dès qu'elle trouverait le moyen de quitter la Russie, ce que d'après les Setlow elle était particulièrement impatiente de faire. »¹²³. Friedberg ajoute : « Plusieurs années après leur rencontre, Waldstein décrocha la permission de quitter l'Union Soviétique, et en profita

¹²² L'original : « Soviet geneticists are apparently divided into two warring camps: those who believe in "creative Darwinism" (akin to Lysenkoism) and those individuals who have been influenced by Western work, such as Dubinin (whose 1956 article shows acquaintance with the work of Benzer, Muller, Hershey and Chase, and the Watson-Crick DNA model). » (Setlow et Setlow, 1961).

¹²³ L'original : « any time that she could manage to leave Russia, which Waldstein had confided to the Setlows she was passionately eager to do. » (Friedberg 1997, p. 94).

pour faire un petit détour par le laboratoire de Setlow en 1974. »¹²⁴ (Friedberg 1997, p. 94). Les données bibliographiques indiquent que Waldstein s'est ensuite installée en Israël, d'où elle a publié une dizaine d'articles avec le mot-clef « *DNA repair* » entre 1974 et 1990.

Une dernière illustration de l'isolement relatif dans lequel se trouvaient les chercheurs soviétiques du point de vue des chercheurs occidentaux, mais de la reconnaissance dont ils disposaient parmi les spécialistes de la réparation de l'ADN, est l'impossibilité dans laquelle s'est trouvée l'Association International de Recherche sur les Radiations (IARR) d'organiser son deuxième congrès en Russie. Au début des années 1960, « d'importants efforts ont été réalisés pour que la communauté russe de recherche sur les radiations s'intègre à l'association, et il a même été envisagé que le second congrès de l'association se tienne en Russie. Cela s'est avéré impossible dans les faits ; et l'Angleterre a été choisie comme pays d'accueil pour l'évènement. »¹²⁵. Lorsque le congrès a eu lieu à Harrogate (Royaume-Uni) en 1963, un chercheur soviétique est tout de même membre du Conseil de l'association. C'est Alexander M. Kuzin (chef de l'équipe de Puschino dans la banlieue de Moscou, que l'on repère sur la Figure 77 avec Ajub I. Gaziev), élu Conseiller pour la Biologie¹²⁶. Kuzin s'est aussi fait connaître sur la scène internationale pour avoir appartenu au mouvement antinucléaire Pugwash¹²⁷, et il continue encore aujourd'hui de co-écrire avec Gaziev depuis Puschino.

Toutes ces histoires confirment l'existence de liens entre les pionniers soviétiques et les pionniers occidentaux malgré la faible liberté de déplacement à la fois durable et « temporaire » des premiers (l'expression est empruntée à Torre, 2008, on y reviendra).

En Europe, outre les associations de recherche comme l'*European Radiation Research Society* (ERRS en lien avec l'IARR), deux organismes soutiennent l'organisation d'un collectif de chercheurs travaillant sur la réparation de l'ADN : il s'agit d'EURATOM et de l'Agence Internationale de l'Energie Atomique (AIEA). EURATOM est constitué en 1957 par les membres de la Communauté Européenne du Charbon et de l'Acier. Cet organisme contribue dès 1973 à financer des recherches collaboratives au niveau européen en faveur de l'utilisation pacifique de l'énergie atomique. De son côté, l'AIEA est un organisme international émanant de l'ONU chargé de veiller aux bons usages de l'énergie atomique.

¹²⁴ L'original : « Several years after this meeting, Waldstein was permitted to emigrate from the Soviet Union and in 1974 she spent a brief stint in Setlow's laboratory. » (Friedberg 1997, p. 94).

¹²⁵ L'original : « Strong efforts were made to involve the Russian radiation research community in the association and it was proposed that the second congress be held in Russia. In the event this did not prove possible; England was chosen as the host country. ». Source : site internet de l'IARR (voir note suivante).

¹²⁶ Source : URL : http://c.ymcdn.com/sites/www.radres.org/resource/resmgr/Docs/IARR_History_Updated_2015.pdf

¹²⁷ Source : URL : http://fr.wikipedia.org/wiki/Mouvement_Pugwash

L'un des centres de recherche de l'AIEA est basé dans la région de Vienne en Autriche d'où publie le groupe d'auteurs autrichiens (Figure 77). Curieusement, on ne trouve aucune trace de ce groupe dans les récits et documents disponibles sur l'histoire de la réparation de l'ADN. Nous pensons que leur activité sur le sujet s'explique par la présence de l'AIEA et par la tenue en 1966 à Vienne d'une rencontre sur les mécanismes de réparation dont se souvient Hanawalt (Hanawalt, 2003, *op. cit.*). Dans les souvenirs d'Hanawalt, cette rencontre a attiré beaucoup moins de monde que celle organisée à Chicago un an plus tôt, mais c'est pour lui l'occasion d'endendre les présentations de deux grands spécialistes européens de radiobiologie : le français R. Devoret et le néerlandais A. Rörsch (*ibid.*). D'après les marqueurs bibliométriques, le groupe autrichien cessera définitivement son activité sur la question au cours des années 1980.

La situation est différente pour les équipes franco-belges et japonaises qui contribuent à la découverte de plusieurs mécanismes de réparation de l'ADN dans les années 1970 (Sur la participation des Japonais à la découverte du BER, voir : Friedberg 1997, 114-115). Paradoxalement, les Français et les Japonais sont volontiers évoqués par les Américains alors qu'ils sont moins visibles que les Autrichiens dans le corpus bibliométrique. Dans les récits, en particulier les récits de pionniers français, on découvre que les Français venaient régulièrement au contact des équipes américaines (notamment à Oak Ridge, Stanford, Yale et Bethesda qui sont centrales sur la Figure 77). Par exemple, Devoret raconte s'être procuré des cellules *E. coli* sensibles auprès de P. Howard-Flanders¹²⁸ à Yale pour les distribuer en Europe et jusqu'en URSS (Devoret, 2001). À propos de la Conférence Internationale des recherches sur les Radiations qui se tient à Évian en 1970, Devoret (Institut du Radium, Paris) se souvient : « Une fois encore, le nombre et la qualité des chercheurs américains participant à un tel événement m'impressionna ; parmi eux, l'un des plus important était Dick Setlow. »¹²⁹ (*ibid.*). Ce témoignage laisse penser que la domination du champ était bien américaine à l'origine. Pour autant, une active communauté de chercheurs européens issus pour la plupart de la radiobiologie et bénéficiant de soutiens financiers européens et internationaux participe

¹²⁸ Paul Howard-Flanders est un radiologiste venu d'Angleterre dans les années 1950 qui contribue avec son post-doctorant Boyce à la mise en évidence du mécanisme de réparation de l'ADN par excision de nucléotide (NER) dans le département de Yale. Tous deux publient des preuves de l'existence de ce mécanisme dans le même numéro qui publie la découverte de Setlow en 1964. D'après ce que nous avons pu comprendre, la discussion sur la priorité de la découverte n'a jamais été entièrement réglée (Yi, 2007). En revanche, davantage qu'Howard-Flanders ou Boyce, Setlow et Hanawalt ont cherché à constituer un collectif et une spécialité autonome de recherche autour des mécanismes de réparation de l'ADN, en témoigne le nombre considérable de publications qu'ils ont consacré à la question.

¹²⁹ L'original : « Once more, the number and quality of the American scientists participating in such an international meeting impressed me. Among them, the most prominent was Dick Setlow, who had discovered nucleotide excision repair. » (Devoret, 2001).

avec les États-Unis à la formation d'un espace mondial pour la communauté au cours des années 1970 et 1980. En effet, il est intéressant de signaler que dès les années 1960, Devoret et Rörsch ont organisé des cours au CEA de Saclay et au TNO de Rijswijk, sur le modèle des cours de Cold Spring Harbour, afin de populariser les recherches biologiques sur les radiations (Friedberg, 1997, *op. cit.* ; Devoret, 2001, *op. cit.*). M. Radman qui est maintenant un scientifique français reconnu pour avoir prouvé l'existence du phénomène « SOS repair » a assisté à ces cours européens de biologie des radiations au début de sa carrière (Friedberg, 1997, 171-83). D'après Friedberg, Radman aurait été envoyé à ces cours par son directeur de thèse Maurice Errera, chef de l'équipe belge de radiobiologie à l'Université Libre de Bruxelles. Après un post-doctorat à Gif-sur-Yvette, Radman est parti pour les États-Unis, et c'est sous la direction de Devoret mais en partenariat avec Evelyn Witkin (visible sur la Figure 77) qu'il a démontré l'existence du système SOS cellulaire — « une réponse cellulaire aux troubles de l'ADN » selon ses propres termes¹³⁰.

Au-delà des cours, et après plusieurs années un peu troubles pour EURATOM, les premiers projets de recherche internationaux sont financés par l'organisme au début des années 1970, parmi lesquels celui d'un consortium d'équipes travaillant sur la réparation de l'ADN. À partir de 1973, ce programme est renouvelé régulièrement jusqu'à aujourd'hui. C'est donc en partie grâce à lui qu'a pu s'établir une collaboration régulière entre les équipes néerlandaises (principalement Leyde et Rotterdam), anglaises (principalement Brighton), françaises (principalement en région parisienne), et italienne (principalement à Pavie) dans le domaine de la réparation de l'ADN. Porteuse d'un certain « internationalisme scientifique », EURATOM a ainsi joué un rôle pionnier dans le soutien des collaborations scientifiques à l'échelle européenne. Il se pourrait même que ces premiers dispositifs aient inspirés les Programmes Cadres de Recherche Européens (PCRD) qui datent des années 1980. De ce point de vue, la spécificité du champ de la réparation de l'ADN, et en particulier son lien au domaine des recherches sur les radiations a été un avantage. De même, le récit de Friedberg sur les conditions d'émergence de la spécialité au Japon (fondé en grande partie sur le témoignage de Bernard Strauss) suggère qu'en raison du manque d'équipements et de fonds alloués à la recherche au sortir de la guerre, il était plus opportun de se lancer dans un nouveau domaine de la biologie comme la réparation de l'ADN, que de chercher à rattraper son retard dans des domaines plus classiques et déjà plus avancés nécessitant davantage de moyens (Friedberg, 114-115, *op. cit.*). Le caractère « déviant » de la spécialité, mais aussi son encastrement dans des champs internationaux comme la radiobiologie, auraient donc pu jouer en sa faveur pour son développement dans l'Europe et le Japon d'après guerre.

¹³⁰ Voir son discours d'entrée à l'Académie des Sciences. URL : http://www.academie-sciences.fr/academie/membre/s170603_radman.pdf

Toutefois, les contextes européen et japonais sont très peu visibles sur la Figure 77, ce qui vient peut-être du fait que le mot-clef « *DNA repair* » a mis plus de temps pour se diffuser dans ces espaces ; et qu'en dehors des États-Unis, il n'y a encore qu'un nombre limité de protagonistes qui en ont connaissance et qui s'y réfèrent au début des années 1970.

Mettre les données bibliométriques en face des récits de pionniers du domaine de recherche nous a permis de ne pas perdre de vue la différence entre l'histoire officielle et l'activité mesurable. Les récits confirment que des recherches sur la réparation du matériel génétique étaient menées dans la plupart des pays disposant de centres de recherche sur l'énergie atomique bien avant l'entrée du mot-clef « *DNA Repair* » à la *National Library of Medicine*. Nos données montrent que les Américains, en particulier Setlow et Hanawalt, ont contribué à populariser l'idée qu'une spécialité de recherche portant sur l'étude des mécanismes de la réparation de l'ADN pouvait se développer. Ils ont eu les moyens et la volonté de promouvoir un collectif de recherche international sur la réparation de l'ADN qui s'est autonomisé peu à peu de la radiobiologie et de la photobiologie pour s'intégrer à la biologie moléculaire. Plusieurs trajectoires de pionniers traduisent cette évolution. La représentation de trajectoires sous la forme d'un graphe permet d'analyser la construction d'un espace d'interactions scientifiques avec ses centres et ses périphéries.

En plus de son apport à la géographie des sciences, cette étude est novatrice à deux égards. D'une part, nous avons analysé la mobilité géographique des scientifiques en s'appuyant sur des données bibliométriques, ce qui est peu fréquent dans le cadre des études de spécialité (Laudel, 2003, 2005 ; Frenken, Hardeman, & Hoekman, 2009). D'autre part, sensible à la nécessité de mieux saisir les phénomènes relationnels en dynamique (Taramasco, Cointet, & Roth, 2010 ; Maisonobe, 2013a), nous avons proposé une utilisation de l'analyse de réseaux à partir de données longitudinales.

2.3. La « communauté scientifique » comme un réseau de lieux

Tout bien considéré, les changements d'affiliation des 35 pionniers retenus ne montrent qu'un aspect des échanges ou circulations ayant structuré l'espace de recherche au stade d'émergence. Cet aspect, celui de la mobilité professionnelle des pionniers, nous en apprend peu sur la mise en place de liaisons internationales, mais suffisamment pour confirmer le rôle central joué par les institutions américaines dans cette histoire.

En fait, les récits et autres marqueurs bibliométriques mobilisés au fil de la section précédente suggèrent que d'autres types d'échanges sont intervenus pour favoriser le développement de la spécialité et son expansion dans plusieurs pays à la fois. En particulier, nous avons évoqué les écoles thématiques, les rencontres et séjours de

recherche ponctuels ainsi que la mise en circulation de certains spécimens comme les cellules *E. coli* sensibles. La prise en compte de ces différents vecteurs relationnels est essentielle pour s'expliquer les quelques cas de collaborations internationales trouvés parmi les publications « *DNA repair* » à plusieurs adresses. Même si ces cas ne résument pas l'ensemble des coopérations formelles qui sont intervenues au stade d'émergence (en raison de la fréquence des données manquantes et du mode d'échantillonnage), ils indiquent que le processus de formation d'un espace communautaire ou « collège » autour du thème de la réparation de l'ADN était déjà bien entamé au milieu des années 1970 (Figure 79).

Collaborations internationales du corpus de publications "DNA Repair" de 1965 à 1975 extrait de la base Scopus

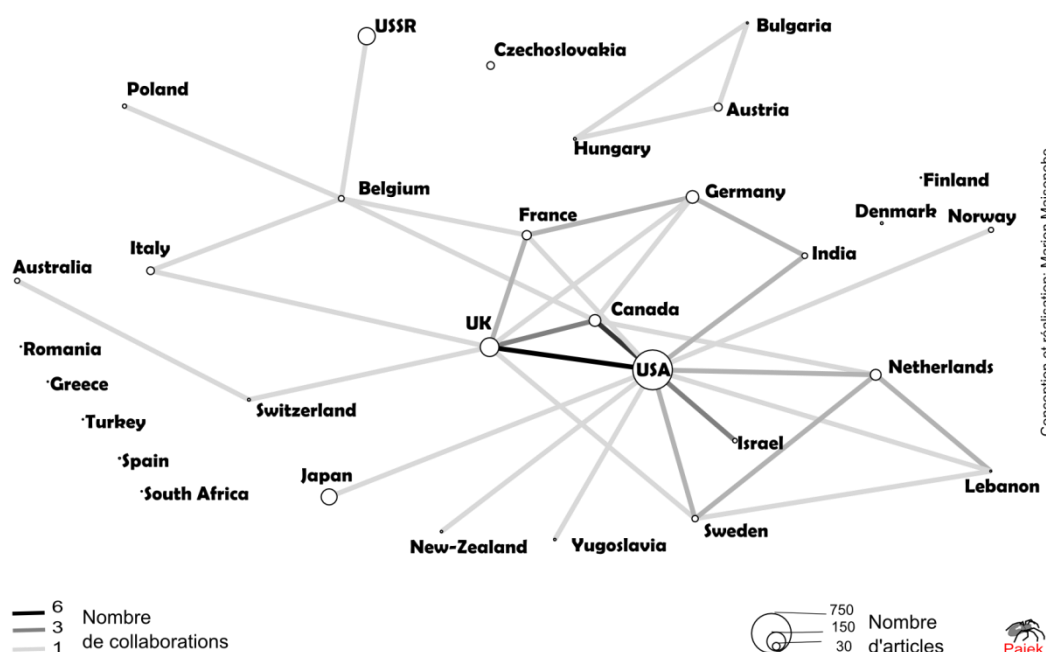


Figure 79 – Connexions internationales d'après les co-signatures enregistrées dans le corpus « *DNA repair* » entre 1965 et 1975, extrait de la base Scopus

Dans l'ensemble, la force des liens représentée est très faible : de 1 à 6 co-signatures d'articles. Néanmoins, la position centrale des États-Unis ressort bien puisque c'est le pays qui a produit le plus d'articles (750 publications au total) et qui a participé au plus grand nombre de collaborations internationales. Ainsi, les chercheurs américains ont signé des articles sur la réparation de l'ADN avec pas moins de 10 pays différents tandis que plusieurs des pays ayant co-publié sur ce thème n'ont eu que les américains pour seuls collaborateurs (en particulier : le Japon, la Norvège et Israël). Trois autres faits saillants doivent être soulignés :

- le triangle de collaborations privilégiées entre les États-Unis, le Royaume-Uni et le Canada ;

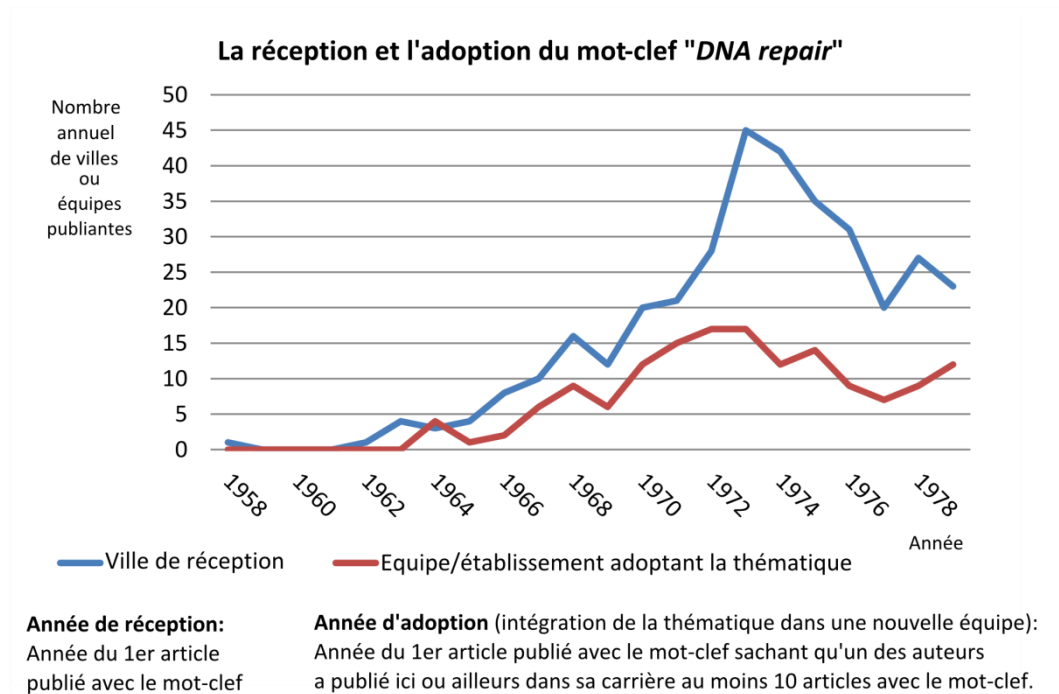
- la position intermédiaire des pays francophones (France, Canada, Belgique) entre l'espace soviétique et l'espace américain ;
- et l'isolement de la composante autrichienne (Autriche, Hongrie, Bulgarie).

Bien que ces observations ne se fondent que sur un très petit ensemble de publications, elles sont révélatrices de l'existence de collaborations internationales formelles sur le thème de la réparation de l'ADN dès les années 1960-1970. Co-signer un article est un acte qui suppose sinon de l'avoir co-écrit, au moins d'avoir apporté une contribution ne serait-ce qu'instrumentale aux résultats qui y sont présentés. Ainsi, en cas de co-signature, on suppose *a minima* qu'il y a eu une interaction préalable entre au moins deux des co-auteurs.

Cette configuration (Figure 79) suggère qu'au moment de la diffusion du mot-clef « *DNA repair* », le collectif en est déjà à un stade de communication ou « réseau » et non plus à un stade paradigmatique. Cette distinction s'appuie sur le travail de Mullins qui décrit plusieurs étapes dans le processus social pouvant conduire à l'institutionnalisation d'une spécialité : paradigme, réseau, cluster, et spécialité (Mullins, 1962). Dans son modèle, le cadre cognitif et le contexte institutionnel caractéristiques d'une spécialité se construisent progressivement dans l'interaction. Il existe une coïncidence entre d'une part, l'évolution de la structure sociale du groupe qui passe d'un stade de faible connectivité à un stade d'échanges intensifs et organisés entre chercheurs et, d'autre part, la progression scientifique elle-même, qu'il analyse suivant un schéma inspiré des travaux de Thomas Khun en distinguant le stade paradigmatique (celui de l'émergence du collectif), et les stades plus avancés, ordinaires et balisés de la pensée scientifique comme celui de la résolution de « puzzles ». Une fois que les problèmes sont bien posés et le cadre de pensée ou paradigme pour les résoudre sont plus stables, alors le collectif entre dans cette phase prolifique sur le plan scientifique, mais moins créative. À ce stade, la dynamique sociale du groupe est à son apogée, le groupe est attractif car prometteur jusqu'à ce que petit à petit les problèmes les plus faciles à résoudre dans le cadre établi soient traités, et qu'il ne reste plus que des problèmes difficiles exigeant une refonte ou du moins une évolution du paradigme en vigueur. Dès lors, en se fondant sur une analyse des relations sociales des membres du « Groupe Phage » (relations de co-signatures, d'apprentissage et de co-appartenance à un laboratoire), Mullins décrit le passage du stade paradigmatique au stade « réseau » comme résultant de la mise en place des premiers échanges et de l'intégration de nombreux membres ; et le passage du stade « réseau » au stade « cluster » comme un effet du renforcement et de la systématisation de certaines relations, de l'adoption d'un certain « style de recherche », mais aussi de la prise de conscience d'appartenir à un collectif organisé. Enfin, le stade de la spécialité est atteint quand la pérennité du groupe ne dépend plus de la participation de tel ou tel membre.

Dans la conclusion de l'étude, on comprend que la plus ou moins grande attractivité des questions de recherche au stade « réseau » détermine le potentiel futur du domaine à faire « spécialité ».

En plus d'être un espace de communication, nos résultats suggèrent que l'espace de recherche sur la réparation de l'ADN est déjà attractif en 1968-1973. Sur la période que nous avons qualifié « d'émergence », le mot-clef « *DNA repair* » se diffuse très rapidement et on compte, en parallèle, une augmentation considérable du nombre d'équipes de recherche qui « adoptent la thématique », c'est-à-dire dont au moins un membre a publié au moins 10 articles sur le sujet dans sa carrière, quelle qu'ait été son affiliation avant cette première publication. Comme le montre la Figure 80, le nombre de nouvelles équipes obéissant au critère d'adoption de la thématique augmente à peu près régulièrement de 1964 à 1973 (17 équipes entrantes en 1973) pour redescendre aux alentours de 5 adoptions par an en 1977. Cette transition dans le rythme d'agrandissement du groupe semble assez bien correspondre à celle que Mullins décrit comme le passage du stade « réseau » au stade « cluster ». Le fait que la coopération de plusieurs équipes européennes sous l'égide d'EURATOM démarre en 1973 (on en voit pas encore la trace sur la Figure 79) pour ensuite être renouvelée régulièrement jusqu'à aujourd'hui est un signe parmi d'autres que la période qui suit la période d'expansion va se caractériser par une plus grande stabilité des échanges entre un noyau dur d'équipes en voie de spécialisation.



Source: Scopus et WoS

Figure 80 – Rythme de réception et d'adoption du mot-clef « *DNA repair* » (1958-1980)

Alors que la « réception » n'implique pas forcément un contact tangible et l'investissement dans une question de recherche directement liée au thème de la réparation de l'ADN, on suppose que l'« adoption » du mot-clef (le fait de le mobiliser fréquemment) est sous-tendue par une véritable intégration au réseau, qui peut se concrétiser à tous moments sous la forme de collaborations formelles.

La réception, c'est-à-dire le simple fait de mentionner le mot-clef dans le titre ou l'*abstract* d'une publication, exige néanmoins d'en avoir eu connaissance. Pour appréhender ce phénomène, on peut considérer la géographie des réseaux de références ou citations. En particulier, les références utilisées par les auteurs rendent compte de leurs influences, et notamment de la littérature scientifique à laquelle ils ont accès et sur laquelle ils s'appuient pour effectuer leurs recherches. Aussi, comme il est possible de suivre la diffusion d'une référence dans le temps et l'espace à travers les marqueurs de citation, il est admis que ce type d'observations a de la valeur en histoire des sciences. Leur exploitation est donc classique dans les études visant à saisir la trajectoire historique d'un objet de recherche. Parmi les plus connues et les plus proches de nos préoccupations se trouve celle de Norman P. Hummon et Patrick Doreian sur le développement de la théorie de l'ADN, qui s'appuie sur une analyse de connectivité dans un réseau de citation à travers le temps (Hummon et Doreian, 1989). En revanche, comme les études bibliométriques centrées sur la diffusion géographique des domaines de recherche sont encore rares, il n'existe pas à notre connaissance de moyens stabilisés pour exploiter les liens de nature géographique entre citations ou références.

Bien qu'il s'agisse d'une approche moins traditionnelle, on peut pourtant tenir compte de l'origine géographique des références. Par exemple, le graphe suivant a été obtenu à partir du codage géographique des références du premier article « *DNA repair* » publié par le pionnier britannique Bryn Bridges en 1966 (Figure 81).

Pour construire ce graphe, nous avons repris et adapté la méthode que Béatrice Milard a développé pour interroger les relations entretenus par les auteurs avec leurs références (Milard, 2010, 2014). Chaque nœud est une référence (le numéro correspond à l'ordre d'apparition de la référence dans le texte) et les références sont liées entre elles lorsqu'elles partagent au moins un auteur en commun. Puisqu'il est ici question d'interroger la géographie des références, nous avons également connecté les références dont les auteurs partagent la même localisation géographique. Enfin, des liens en pointillé ont été tracés : d'une part entre la référence signée depuis l'*Institute of Cancer Research* (ICR de Londres) et celle signée depuis le centre de recherche nucléaire d'Harwell qui partagent toutes deux une proximité géographique ; et d'autre part entre les références signées par les auteurs de Chicago, Stanford, Yale, et Oak-Ridge parce qu'on a noté, par ailleurs, qu'ils ont circulé entre ces différentes institutions.

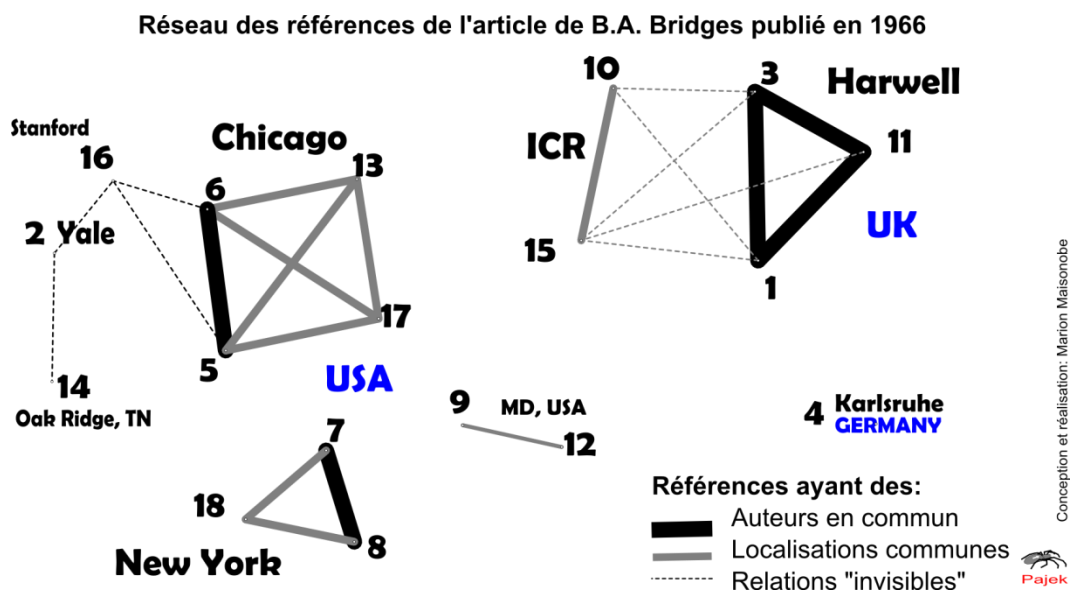


Figure 81 – Réseau des références de l'article « DNA repair » de B.A. Bridges (paru en 1966)

Le résultat montre que lorsqu'il se réfère pour la première fois au mot-clef « DNA repair », Bridges (alors à Harwell) a bien conscience du contexte d'origine de cette désignation et qu'il a pris connaissance du travail des premiers auteurs à l'avoir proposée. En dehors des références à son propre travail et à celui de ses voisins de l'ICR, il ne se réfère qu'à des pionniers américains. La seule exception est une référence à un article allemand qui ne concerne pas directement le thème de la réparation de l'ADN.

Malgré tout le sel de cette représentation, elle ne suffit pas à comprendre comment Bridges, une fois à la tête de l'équipe du MRC de Brighton, deviendra l'un des principaux pionniers européens du domaine (Walker, 2001). Le fait que, comme l'ont montré les données de mobilités professionnelles, le jeune Sedgwick ait été envoyé par Bridges en séjour d'étude aux États-Unis tout comme la fréquence des co-signatures entre auteurs américains et britanniques (Figure 79) sont beaucoup plus révélateurs de l'existence de liens tangibles entre l'espace d'origine du « mot-clef » et la Grande-Bretagne.

L'hypothèse à laquelle on se réfère ici implicitement est celle, très populaire en géographie de l'innovation, de l'importance de la transmission du savoir « tacite » dans l'acquisition de savoir-faire et la mise en place d'une communauté. Le savoir « tacite » désigne le savoir qui n'est pas formalisable, voire même dicible, et donc par définition, celui qui ne peut pas circuler par l'intermédiaire de la littérature. Michael Polanyi est l'auteur à qui cette théorie est généralement attribuée. L'idée est que la manière de faire se transmet dans l'interaction directe, que l'on qualifie parfois de « face à face ».

Cette théorie est souvent mobilisée pour justifier les regroupements géographiques d'entreprises ou d'acteurs au sein de « clusters » (voir Chapitre 2 et Chapitre 3). Reconnu pour avoir forgé le concept de « communauté scientifique internationale », Polanyi considérait comme beaucoup de ses contemporains que la science (« moderne ») était venue d'Europe pour se diffuser dans le reste du monde (Chapitre 4). Ainsi, trouve-t-on une trace de la théorie dans l'interprétation qu'il donne de ce phénomène de diffusion : « Alors que le contenu de la science est enseigné avec succès aux quatre coins du monde dans des centaines de nouvelles universités, l'art de la recherche scientifique n'a toujours pas pénétré en beaucoup d'entre elles. Les régions d'Europe dans lesquelles la méthode scientifique a germé pour la première fois il y a 400 ans sont toujours très fertiles, malgré leur appauvrissement, comparé à des tas de territoires outre-mer disposant de bien plus d'argent pour leurs recherches. Sans l'opportunité offerte à des jeunes scientifiques de suivre un apprentissage en Europe, et sans la migration de chercheurs européens vers les nouveaux pays, les centres de recherche outre-mer pourraient difficilement avoir fait grand-chose jusqu'ici. »¹³¹ (Polanyi, 1958, p. 55).

On trouve là une distinction entre le « contenu » de la science et « l'art de la recherche scientifique », le premier pouvant voyager sans difficulté contrairement au second. Surtout, il est suggéré que la transmission de savoir-faire par-delà les frontières dépend des opportunités de mobilité internationale. Polanyi évoque deux types de déplacements différents qu'il ne cherche pas à différencier : d'une part, le séjour de recherche des jeunes chercheurs à l'étranger ; et d'autre part, la migration des chercheurs. Dans la section précédente, sur les 3 cas de déplacements internationaux repérés à l'aide des marqueurs bibliométriques, des exemples des deux types ont bien été mis en évidence : d'une part, les séjours de recherches post-doctoraux de Sedgwick et Paterson, et d'autre part la migration de Walter Harm, qui peut s'expliquer par le contexte défavorable sur le plan économique et professionnel au sortir de la guerre en Allemagne, et la présence de collègues et d'opportunités de travail aux États-Unis.

L'idée que ces déplacements ont joué un rôle dans la circulation du savoir et dans la structuration du champ des recherches sur la réparation de l'ADN est évidente et justifie qu'on s'y intéresse, tout en gardant en tête que l'approche bibliométrique n'est qu'un moyen parmi d'autres de capturer des informations sur la mobilité des scientifiques.

¹³¹ L'original : « while the articulate contents of science are successfully taught all over the world in hundreds of new universities, the unspecifiable art of scientific research has not yet penetrated to many of these. The regions of Europe in which the scientific method first originated 400 years ago are scientifically still more fruitful today, in spite of their impoverishment, than several overseas areas where much more money is available for scientific research. Without the opportunity offered to young scientists to serve an apprenticeship in Europe, and without the migration of European scientists to the new countries, research centres overseas could hardly ever have made much headway. » (Polanyi, 1958, p. 55).

Surtout, cette approche n'est pas adéquate pour repérer les déplacements courts (tels que les « séjours d'apprentissage ») quand bien même ce type de déplacements participe également à structurer les collectifs sociaux. On l'a déjà évoqué, et l'émergence de la réparation de l'ADN en témoigne, il semblerait que la « proximité temporaire » suffise à activer ou entretenir des relations de coopérations, mais les travaux qui parviennent explicitement à quantifier ce phénomène sont encore rares. Pour y remédier, Bastien Bernela poursuit actuellement, avec Béatrice Milard, des travaux visant à expliquer les collaborations d'une population de mathématiciens et de chimistes français en distinguant différents modes de mises en relation¹³². En complétant les trajectoires bibliométriques de 10 mathématiciens d'un laboratoire donné par leur *curriculum vitae* et des entretiens, ils montrent que les collaborations formelles de ces derniers avec des chercheurs localisés dans la même ville qu'eux ou dans une ville qu'ils ont fréquentée par le passé représentent moins de la moitié de leurs collaborations (36 %). Les autres collaborations sont sous-tendues par des rencontres à l'occasion d'événements scientifiques, de courts séjours de recherche, d'invitations et de participation à des jurys de thèse, et une seule repose sur un contact qui n'a jamais été rencontré. Même si cela reste exceptionnel, il existe effectivement des liens de co-signature entre des chercheurs qui ne se sont jamais rencontrés mais ont communiqué à distance, ce qui peut venir contredire la nécessité même de se déplacer pour collaborer.

Ce dernier cas, nous en avons trouvé un exemple parmi les spécialistes de la réparation de l'ADN interrogés aux Pays-Bas. Dans ce cas, évoqué dans le Chapitre 6, la co-signature formalisait le fait que le chercheur avait contribué matériellement à la publication en envoyant le modèle de souris mutante (dont il est le créateur) à une équipe « inconnue » souhaitant l'utiliser pour ses propres expériences. À ce moment-là, le lien qui sous-tend la collaboration n'est ni un lien de co-location (proximité durable) entre coéquipiers, ni un lien associé à des situations de « proximité temporaire » entre collaborateurs, mais un lien social indirect matérialisé par la circulation d'un objet « non-humain » entre deux lieux. Cet exemple nous amène, après avoir souligné le rôle des contacts humains dans la structuration d'un réseau scientifique d'échelle mondiale, à souligner le rôle de l'instrumentation, des spécimens ou plus généralement des « ressources » nécessaires à la réalisation des travaux de recherche du collectif social étudié sur sa géographie.

¹³² Ces travaux ont fait l'objet de deux récentes présentations : Bernela B. et Milard B., 2015. *Collaborative trajectories of researchers or How publication is an event in a scientific career*. Journées ResTo: « Art, réseaux et trajectoires », 31 Mars-1^{er} avril 2015, Université Toulouse Jean-Jaurès ; Bernela B. et Milard B., 2015. Mobilité géographiques, mises en relations et carrière scientifiques. Journées d'étude du programme ANR Resocit : « Citations scientifiques et réseaux sociaux », 4-5 juin 2015, Université Toulouse Jean-Jaurès

Même si elles ne laissent pas forcément de traces écrites dans les publications, les conditions d'accès aux ressources sont déterminantes pour la géographie du domaine de recherche et son potentiel d'évolution.

Partant de ce constat, il existe des travaux qui se centrent spécifiquement sur la circulation de certaines ressources. En particulier, Angela Creager, spécialisée dans l'histoire des recherches nucléaires de l'immédiat après-guerre a étudié l'exportation de radioisotopes depuis les États-Unis, et plus spécifiquement depuis Oak Ridge, vers différents laboratoires étrangers à des fins biomédicales (Creager, 2002 dont est extraite la Figure 82).

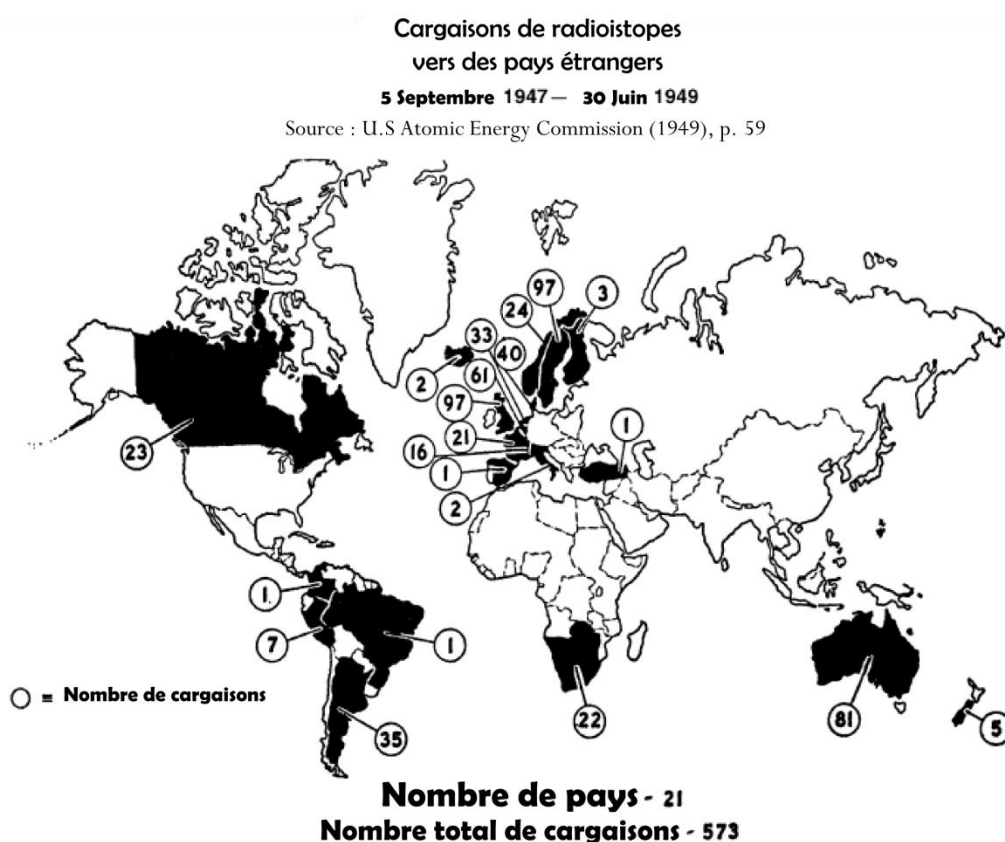


Figure 82 – Carte indiquant la destination des cargaisons de radioisotopes de la Commission de l'Énergie Atomique des États-Unis lors des 22 premiers mois du programme. Extraite et traduite de l'article « Tracing the politics of changing postwar research practices: the export of 'American' radioisotopes to European biologists » (Creager, 2002)

Si la répartition de ces objets particuliers a pu intervenir dans la structuration de la biomédecine et plus précisément de la biologie moléculaire, nous pouvons très bien nous interroger sur leur rôle, à un moment donné, dans le développement des études sur la réparation de l'ADN.

Plus généralement, puisque le développement de la réparation de l'ADN est associé au développement de la biologie moléculaire, on gagnerait sans doute à mettre en regard la géographie de l'émergence de la biologie moléculaire et celle des recherches sur la réparation de l'ADN, puis de repérer le moment où les deux géographies ont commencé à se rejoindre. En effet, comme le note Creager dans la recension qu'elle fait de l'ouvrage de Friedberg, les deux géographies présentent l'originalité d'avoir été dissociées à l'origine : « Une fois que la réparation de l'ADN a été étiquetée comme telle (...), le champ n'a pas été dominé par les chercheurs des centres de biologie moléculaire comme Caltech ou Cambridge : des contributions clefs sont venues d'Oak Ridge, Brookhaven, et plus tard de la "Texas Mafia", qui ont formé un réseau de centres importants pour les recherches sur l'ADN. »¹³³ (Creager, 1999).

De ce fait, situer l'émergence de la réparation de l'ADN par rapport à celle de la biologie moléculaire est un défi (que nous n'avons pas relevé ici, mais) qui, en dehors de son intérêt sur le plan de la géographie des sciences, serait en mesure d'enrichir l'historiographie de la biomédecine. En effet, la marginalité ressentie par les pionniers de la réparation de l'ADN ne se limite pas à leur isolement des débuts, mais se traduit encore aujourd'hui par un certain penchant de l'histoire des sciences pour l'histoire de la découverte des mécanismes de réplication de l'ADN (la reproduction du matériel génétique), au détriment de celle des mécanismes de réparation. La façon, par exemple, dont Michel Morange raconte la découverte d'A. Kornberg, biochimiste de Stanford, nous semble assez symptomatique de ce phénomène : « Les expériences ultimes devaient confirmer que l'enzyme purifiée par Kornberg n'était pas l'enzyme essentielle pour la réplication de l'ADN mais intervenait dans la réparation de celui-ci. Les enzymes participant à la réplication *in vivo* de l'ADN avaient cependant une structure proche de l'ADN polymérase purifiée par A. Kornberg et surtout fonctionnaient suivant les mêmes principes. La découverte du rôle physiologique mineur de l'ADN polymérase de Kornberg ne remettait pas en cause la démonstration que les enzymes étaient capables de répliquer l'ADN conformément au modèle de Watson et Crick. » (Morange, 2003, p. 312). Ce qui importe dans cette description n'est pas l'intérêt de cette découverte pour la compréhension de la réparation de l'ADN, mais son intérêt, même indirect, pour celle de la réplication. Or cette découverte, qui a eu lieu à Stanford, a été évidemment très importante pour le développement des recherches sur la réparation de l'ADN.

¹³³ L'original : « Once DNA repair had been labeled as such, [...] neither was the field dominated by workers from molecular centers such as Caltech and Cambridge [End Page 541] University: key contributions came from Oak Ridge and Brookhaven National Laboratories, and subsequently from the "Texas mafia," as a network of important centers for DNA came to be known. » (Creager, 1999).

En fin de compte, la géographie des recherches sur la réparation de l'ADN dépend à l'origine, pour ce que nous avons pu en voir, de la géographie des recherches sur les radiations, puis s'est associée petit à petit à la géographie de la biologie moléculaire, mais aussi à celle des recherches sur le cancer et le vieillissement (on peut parler d'« encastrement »). Ces effets d'encastrement s'expliquent par la répartition ou l'accès à des ressources nécessaires aux pratiques de recherche. La localisation d'une équipe de recherche est ainsi sensible à différents facteurs tels que : l'ouverture de nouveaux départements de recherche à l'université, la répartition des instruments techniques et des savoir-faire, la répartition de certains équipements lourds, des budgets etc. mais aussi à l'accès par le biais des circulations et des échanges à des idées, de la main d'œuvre, et du matériel. Il faut aussi penser que les besoins varient en fonction des questions de recherche, ce qui peut expliquer les évolutions dans la géographie observée. Ainsi, comme on l'a évoqué en début de chapitre, le domaine de la réparation de l'ADN a très tôt été constitué de plusieurs sous-ensembles de recherche. Dès le départ, on pouvait déjà noter que les localisations et les échanges variaient en fonction du type d'organisme étudié (mammifère ou non), puis à mesure que des mécanismes distincts de réparation de l'ADN ont été mis en évidence, des travaux spécifiques ont été menés sur chaque mécanisme.

Dans le chapitre suivant, la focale est justement placée sur l'un de ces sous-ensembles de recherche, et plus précisément sur les travaux associés à une question précise apparue dans les années 1990, qui a mobilisé des spécialistes de la réparation de l'ADN par excision de nucléotide (NER). Ce mécanisme est celui, déjà évoqué, dont la découverte en 1964 a généré la promulgation et l'adoption du mot-clef « réparation de l'ADN ». Ensuite, en 1968, James Cleaver a démontré que les enfants de la lune, atteints de *Xeroderma Pigmentosum*, étaient porteurs de cellules incapables de réparer les lésions causées par les rayonnements solaires. Une fois avoir compris qu'une défaillance de leur système de réparation de l'ADN était responsable de leur hyper-sensibilité aux rayons UV, de nombreux travaux ont été entamés s'appuyant sur l'analyse de cellules d'enfants atteints par la maladie orpheline, dites cellules XP (pour *Xeroderma Pigmentosum*). Dans les années 1990, une partie des chercheurs de cette branche interrogent les liens existants entre le NER et le mécanisme de transcription de l'ADN. La traduction spatiale de cette dynamique ou émulation scientifique est l'objet du dernier chapitre.

CHAPITRE 11. LA DIFFUSION SPATIALE D'UNE QUESTION DE RECHERCHE

Comment évolue la géographie d'un collectif lorsqu'il s'agit de résoudre une question de recherche ? Appartenir à une spécialité indique une inclination à se pencher sur un certain type de problèmes, mais le travail particulier à la résolution d'une question de recherche est une affaire plus concrète qui nécessite, cette fois-ci, très directement l'accès à des ressources cognitives, humaines et matérielles voire la mise en place d'un protocole partagé entre les différentes équipes abordant la question. On se concentre ici sur une question en relation avec le mécanisme de réparation de l'ADN par excision de nucléotide (*Nucleotide Excision Repair* dit « NER »). Ce mécanisme constitue un point de départ parmi d'autres autour duquel la spécialité s'est structurée à partir des années 1960. L'étude d'une question spécifique de recherche portant sur le NER dans les années 1990 permet de suivre le développement de la spécialité à travers une de ces ramifications. Ce développement est porté par l'équipe de biologie cellulaire de Rotterdam qui, dès sa création dans les années 1970, a entamé des recherches sur le NER à partir de cellules d'enfants atteints de *Xeroderma Pigmentosum* (XP). Le choix de travailler précisément sur cette maladie a alors été justifié par la découverte de l'américain James Cleaver en 1968, et facilité par la présence, à proximité du laboratoire de D. Bootsma, d'un département hospitalier de dermatologie traitant des patients atteints par la maladie orpheline.

Après avoir étudié la géographie d'un corpus de publications représentatif de la question de recherche pour saisir sa dynamique et son organisation spatiale, l'« univers de référence » d'une chercheuse travaillant actuellement sur le NER et ayant investi précisément cette question de recherche lors de son post-doctorat à Rotterdam permet de comprendre d'encore un peu plus près comment les pratiques de recherche se déploient dans l'espace géographique et participent à sa structuration. À l'issue de ce chapitre, l'ambition est d'arriver, enfin, à concevoir l'espace de la spécialité et sa dynamique comme résultant de l'articulation des différents niveaux d'emboîtements constitutifs de la spécialité de recherche : de celui de la trajectoire individuel à celui de la spécialité en passant par celui de la question de recherche et celui du sous-groupe (ici le NER). Ce travail ouvre aussi des questions sur la position d'une sous-question de recherche dans un ensemble devenu très dense (du point de vue des chercheurs). En effet, on verra s'opérer au niveau individuel une distinction entre réseau de collaboration et réseau de références, le second mettant plus volontiers en avant des liens uniquement cognitifs.

1. Géographie de l'étude du lien entre la transcription et la réparation de l'ADN

« Et puisque le monde intellectuel, royaume des symboles abstraits universellement applicables, transcende les divisions du monde matériel sur lesquelles sont fondées les nations, il est inévitable qu'une étroite allégeance au premier de ces mondes aille de pair avec une tendance à négliger le second. » (Storer, 1970).

1.1. L'expansion géographique de la question de recherche

Afin d'échapper aux effets de taille et d'avoir une meilleure compréhension de l'organisation du travail scientifique dans l'espace géographique, il est intéressant de se rapprocher du contenu des recherches, par exemple en se concentrant sur un corpus de publications bien délimité. Comme on l'a vu en détail dans le chapitre précédent, la démarche consistant à se rapprocher des contenus pour saisir l'organisation des collectifs scientifiques est assez répandue en *science studies*, mais plus rare sont les cas où elle s'accompagne d'une prise en compte de l'inscription spatiale des activités scientifiques. Dans ce chapitre, on s'intéresse au niveau auquel la science se fait, celui de la question de recherche ou *problem area*. Pour Diana Crane, « Une question de recherche est un groupement éphémère de chercheurs qui traitent de problèmes spécifiques jusqu'à ce qu'une ou plusieurs décennies plus tard, les problèmes soient résolus ou déclarés insolubles. »¹³⁴ (Crane, 1969b). À ce niveau, l'organisation sociale, s'il y en a une, est insaisissable, peu structurée et éphémère. Par conséquent, on s'attend à ce que les liens se nouent et se dénouent rapidement dans le réseau de co-publications que nous allons étudier. Nos résultats devraient alors différer, ou du moins faire apparaître de subtiles nuances, par rapport aux résultats obtenus aux niveaux plus agrégés envisagés précédemment : le niveau de la spécialité scientifique, et celui de l'activité de production scientifique en général.

L'analyse porte sur une question de recherche en biologie moléculaire qui présente des caractéristiques semblables à l'objet *problem area* décrit par Diana Crane. Il s'agit d'une thématique de recherche qui regroupe, à partir des années 1990, des scientifiques tâchant de comprendre le lien entre réparation et transcription de l'ADN. Dans un premier

¹³⁴ L'original : « The problem area can best be understood as a temporary unit which deals with special problems and then dissolves after one or several decades when the problems have either been solved or been determined to be unsolvable. » (Crane, 1969b).

temps, on décrit la construction d'un corpus bibliographique représentatif de la question de recherche. La diffusion spatiale des publications est ensuite analysée, puis l'évolution du réseau de co-publications ; ce qui permet d'interroger non plus seulement la propagation du collectif scientifique dans l'espace géographique, mais aussi l'inscription spatiale de sa structure. Il ne s'agit pas seulement de comprendre les mécanismes d'évolution du collectif mais aussi de se demander si ces mécanismes peuvent être influencés par des logiques spatiales.

Pour extraire un corpus de publications représentatif de la question de recherche, on procède à une extraction de données depuis le *Web of Science*. L'extraction est réalisée à partir d'une requête par mots-clefs. L'ensemble de mots, devant être retrouvés dans le titre, le résumé et les mots-clefs des notices de la base, sont sélectionnés en concertation avec une biologiste moléculaire travaillant sur la question de recherche considérée. Il s'agit du nom de trois maladies et des termes « transcription » et « réparation » de l'« ADN » en anglais¹³⁵. Les enfants atteints de ces maladies orphelines (dont la Maladie de la Lune) souffrent d'une déficience dans la réparation des lésions produites sur l'ADN par les UV. En 1993, l'équipe spécialisée dans la réparation de l'ADN à Rotterdam a mis en évidence, avec une équipe de biochimistes strasbourgeoise, spécialisée dans la transcription de l'ADN, l'existence d'un lien entre la réparation des lésions dues aux UV (ou réparation par Excision de Nucléotide) et le facteur de transcription TFIIH. La publication dans la revue *Science* du résultat obtenu par ces deux équipes reste à ce jour l'article le plus cité du corpus de publications (avec 662 citations reçues au 15 Juin 2015, voir Figure 83). À la suite de cette découverte, les travaux tâchant de caractériser le lien unissant réparation et transcription en se concentrant sur l'une ou plusieurs de ces trois maladies orphelines se sont multipliés. Il faut dire que la découverte a bénéficié d'une bonne « médiatisation » par le biais de la revue *Science*, d'autant que l'année suivante le domaine de la réparation de l'ADN a été à l'honneur en recevant la distinction « *Molecule of the year* » délivrée par le rédacteur en chef de la revue (Koshland, 1994). Ce dernier reconnaît enfin l'importance des mécanismes de la réparation de l'ADN, ce qui symboliquement est important pour les membres de cette spécialité, longtemps restée aux marges de la biologie moléculaire : « Si l'ADN était mal copié, on attraperait des maladies comme le cancer beaucoup plus souvent, et nous n'aurions pas de copie fidèle de notre héritage parental. Notre espèce ne serait pas préservée, et on ne vivrait pas longtemps. Si l'ADN était parfaitement copié, il n'y aurait pas de place pour l'évolution, et le principe de l'apparition de nouvelles espèces

¹³⁵ La requête effectuée sur la base de données du WoS a été la suivante : ("xeroderma pigmentosum" OR "cockayne syndrome " OR "trichothiodystrophy") AND ("DNA" AND "repair" AND "transcription").

mieux adaptées à l'environnement aurait disparu depuis bien longtemps. Le système de réparation de l'ADN offre un bon compromis. »¹³⁶ (*ibid.*).

DNA-REPAIR HELICASE - A COMPONENT OF BTF2 (TFIIH) BASIC TRANSCRIPTION FACTOR

By: SCHAEFFER, L (SCHAEFFER, L); ROY, R (ROY, R); HUMBERT, S (HUMBERT, S); MONCOLLIN, V (MONCOLLIN, V); VERMEULEN, W (VERMEULEN, W); HOEIJMAKERS, JHJ (HOEIJMAKERS, JHJ); CHAMBON, P (CHAMBON, P); EGLY, JM (EGLY, JM)

SCIENCE

Volume: 260 Issue: 5104 Pages: 58-63

DOI: 10.1126/science.8465201

Published: APR 2 1993

[View Journal Information](#)

Abstract

The human BTF2 basic transcription factor (also called TFIIH), which is similar to the delta factor in rat and factor b in yeast, is required for class II gene transcription. A strand displacement assay was used to show that highly purified preparation of BTF2 had an adenosine triphosphate-dependent DNA helicase activity, in addition to the previously characterized carboxyl-terminal domain kinase activity. Amino acid sequence analysis of the tryptic digest generated from the 89-kilodalton subunit of BTF2 indicated that this polypeptide corresponded to the ERCC-3 gene product, a presumed helicase implicated in the human DNA excision repair disorders xeroderma pigmentosum and Cockayne's syndrome. These findings suggest that transcription and nucleotide excision repair may share common factors and hence may be considered to be functionally related.

Keywords

KeyWords Plus: RNA-POLYMERASE-II; PUTATIVE HELICASES; INITIATION-FACTOR; REPLICATION; PHOSPHORYLATION; PROTEINS; PROMOTER; REQUIREMENT; COMPLEXES; CLONING

Author Information

Addresses:

[1] FAC MED STRASBOURG,CNRS,UPR 6520,INSERM,U184,11 RUE HUMANN,F-67085 STRASBOURG,FRANCE

[2] ERASMUS UNIV,CTR MED GENET,DEPT CELL BIOL & GENET,3000 DR ROTTERDAM,NETHERLANDS

Figure 83 — Notice de publication de l'article démontrant le lien entre réparation et transcription de l'ADN. Source : Web of Science

Les premières publications du corpus paraissent à partir de 1991 et leur nombre décolle après 1993 (courbe rouge, Figure 84). Pour une analyse longitudinale, le corpus a été découpé en deux périodes de même taille (7/8 ans)¹³⁷ : 1991-1999 et 2000-2007. La première période est une phase d'émergence et d'émulation scientifique. 40 % des publications sont publiées dans la première période. Le nombre de publications progresse jusqu'à atteindre son maximum en 2001 (au début de la période suivante). Il chute et stagne (aux alentours de 50 articles publiés par an) au cours de la seconde période, ce qui laisse supposer que la question de recherche tend à s'épuiser.

¹³⁶ L'original : « If the DNA were copied badly, we would have diseases such as cancer at a much higher frequency, and we would not get a faithful copy of our parental inheritance. Our species would not be preserved, and we would not live long. If the DNA were copied perfectly, there would be no room for evolution, and the basis for creation of new species with better environmental adaptation would have vanished long ago. The DNA repair system allows a happy medium. » (Koshland, 1994).

¹³⁷ Pour que les deux périodes soient statistiquement comparables, nous avons fait le choix qu'elles soient de même taille. Le choix de se contenter de deux périodes est cohérent compte tenu de la taille de l'échantillon étudié (moins de mille publications au total).

Comme l'a montré Mullins à propos du Groupe Phage, le passage d'un stade de recherche à un autre est une clef d'interprétation indispensable pour interpréter l'évolution du groupe scientifique. Mais à la différence de l'objet étudié par Mullins, l'objet de type *problem area*, étudié ici, n'existe que par rapport à une seule question de recherche. Les protagonistes ne se définissent pas par leur appartenance à cette question de recherche, mais par leur appartenance à des communautés scientifiques plus vastes. En l'occurrence, la question étudiée est investie en majorité par des membres de la communauté de la réparation de l'ADN (et plus spécifiquement par le groupe NER pour « *Nucleotide Excision Repair* ») et dans une moindre mesure par des membres de la communauté de la transcription de l'ADN (dont l'équipe strasbourgeoise). Le collectif étudié est donc hybride. Il n'est pas institutionnalisé et on imagine qu'il peut se dissoudre rapidement. Ce type d'évolution est connu mais la traduction spatiale d'un tel processus l'est beaucoup moins.

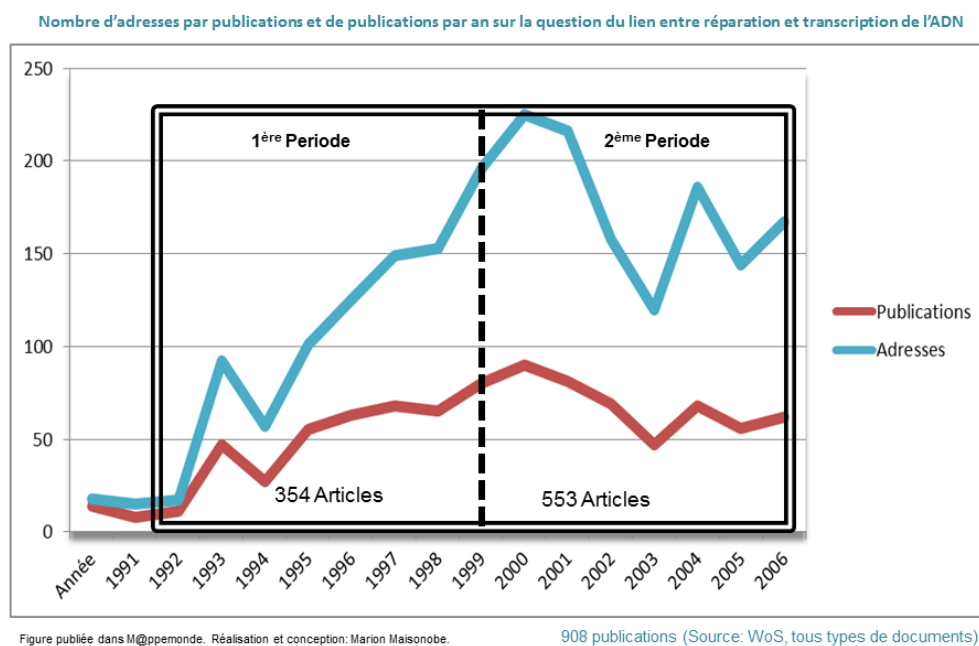


Figure 84 – Corpus de publications extraites du Web of Science à l'aide de mots-clefs sur le lien entre réparation et transcription de l'ADN : sélection de deux périodes d'étude

Pour cette étude, plusieurs niveaux géographiques d'agrégation des données ont été considérés pour finalement privilégier celui très englobant de l'agglomération. En agrégeant l'information au niveau des « agglomérations scientifiques », le corpus étudié est passé de 298 localités à 229 agglomérations. Comme on l'a montré dans la Partie 2.Chapitre 7.2.2, le choix du découpage en agglomérations n'a pas été sans effet sur les observations tirées de ce corpus, particulièrement dans les cas où comme pour la baie de

San Francisco, le choix d'un regroupement plutôt qu'un autre modifiait l'allure du réseau de co-publications.

Dans le corpus, la fréquence des collaborations augmente progressivement. Les premiers articles signés par plusieurs agglomérations à la fois sont publiés en 1991, mais les premières collaborations internationales en 1993. Comme la Figure 84 (courbe bleue) le montre, le nombre d'adresses par article augmente jusque dans les années 2000 (de 2 à 2,5 adresses par article en moyenne) ce qui, en raison de la densification de l'activité intra-locale, se traduit par une augmentation plus limitée du nombre moyen d'agglomérations distinctes ayant co-signé un article (de 1,65 à 1,76).

Après avoir décrit l'expansion géographique de la question de recherche, l'analyse se concentrera plus précisément sur les logiques spatiales d'évolution du réseau de collaborations scientifiques entre agglomérations investies dans la question de recherche.

Pour saisir l'évolution de la répartition de l'activité scientifique portant sur cette question, on s'appuie d'abord sur la Figure 85 qui représente les agglomérations actives dans le corpus : en gris-bleu, les quelques agglomérations qui ne sont actives qu'au cours de la première période (32 agglomérations) ; en violet, les 87 agglomérations qui publient à la fois dans la première et la seconde période, et en fuchsia, celles qui ne publient que pendant la seconde période (elles sont 110).

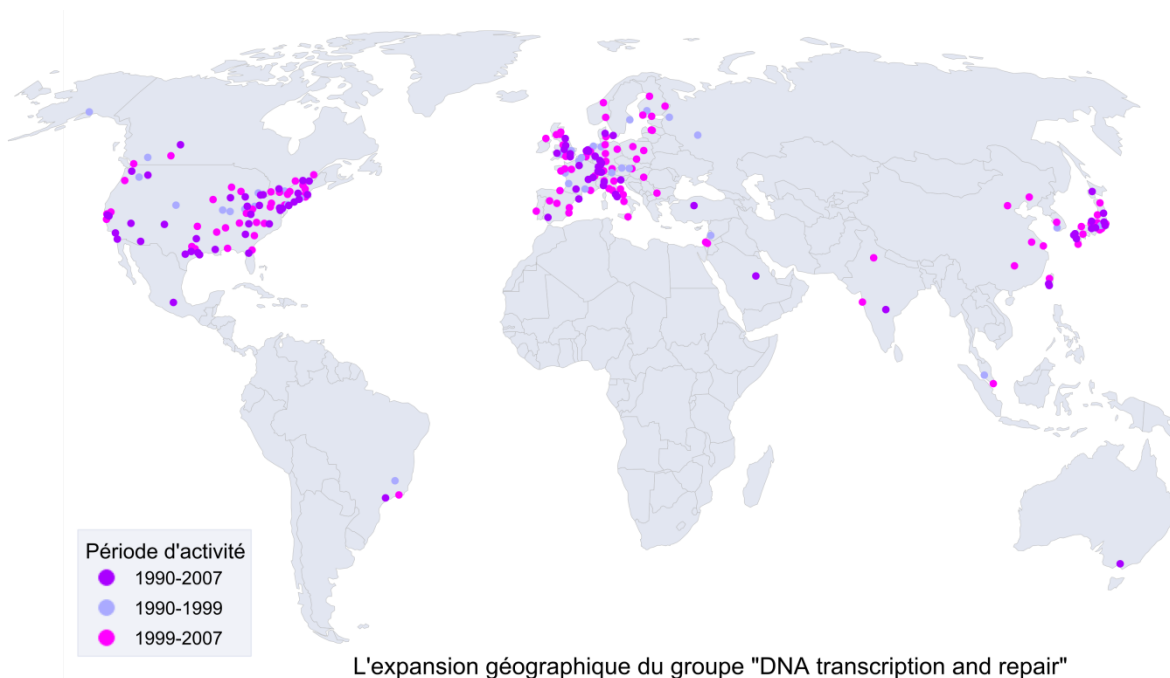


Figure 85 – La répartition des agglomérations publiantes selon l'année de leur première publication relative à la question de recherche. Source : Web of Science

Les plus fortes densités d'agglomérations publiantes se trouvent autour des premiers laboratoires à avoir publié sur la question de recherche : aux États-Unis (sur la Côte Est américaine, en Californie et au Texas) ; en Europe, autour du consortium à l'origine de la question de recherche (Leyde-Rotterdam-Strasbourg) ; et au Japon.

Près de 80 % des 110 nouvelles agglomérations se trouvent dans seulement 10 pays (dans l'ordre : USA, Japon, Italie, Royaume-Uni, Allemagne, Chine, Espagne, Canada, Finlande, France). Parmi eux, il n'y a que la Chine qui ne comptait aucune agglomération travaillant sur la question au cours de la première période. À cette exception près, la diffusion se fait principalement par « contagion » au sein des pays les plus productifs de la première période (Saint-Julien, 2004). Parmi ces dix pays qui concentrent la majeure partie du phénomène de diffusion, l'Italie, l'Espagne et la Finlande comptent davantage d'entrants dans la seconde période qu'ils n'en avaient dans la première. Cette information est importante car elle montre que la question est suffisamment porteuse ou, du moins, que les résultats scientifiques obtenus sont assez fondamentaux pour justifier un « rattrapage » de certains pays dans la seconde période. Dans la seconde période, la question de recherche s'étend à onze nouveaux pays dont cinq en Europe de l'Est (la Pologne, la Lettonie, la Hongrie, la Bulgarie et la Slovaquie). Les autres entrants sont la Chine, Israël, la Norvège, la Belgique, le Portugal et Singapour. En tout, un nombre relativement limité de pays (36) apparaît concerné par cette question de recherche ce qui vient notamment du degré de spécialisation requis dans le domaine de la réparation de l'ADN pour être en mesure de l'aborder.

Dans ce corpus, la majeure partie des publications est en fait produite par un nombre restreint d'agglomérations : elles suivent la loi de Lotka (Encadré 2). Comme cela était attendu en raison de la nature de l'objet *problem area* : à côté d'un noyau dur d'équipes réellement investies dans la question de recherche (qui publient régulièrement) se trouvent quantité d'équipes dont la participation aux travaux en rapport avec cette question de recherche reste épisodique (un ou deux articles). Dès lors, sachant que peu d'équipes investissent véritablement cette question, on peut distinguer plusieurs catégories d'adoptants à la manière de Everett Rogers (1962), « les innovateurs et primo-adoptants » et les « adoptants tardifs » (la majorité).

Une manière de formaliser cette distinction a été de noter systématiquement l'année du premier article publié pour chaque agglomération puis le nombre de fois (en nombre d'années) où l'agglomération produit à nouveau sur le sujet. Quand ce dernier nombre est rapporté au « nombre potentiel d'années », c'est-à-dire, au nombre d'années restantes (dans la période considérée) à partir de la sortie du premier article, on obtient un indice d'« implication » de l'agglomération. Cet indice est délicat à utiliser car il est extrêmement sensible au temps : plus l'entrée est tardive dans la période considérée,

moins l'indice est significatif. 32 agglomérations ont un indice supérieur à 0,5 dans la première période, c'est-à-dire qu'elles sont présentes (au cours de la première période) plus de la moitié du temps espéré compte tenu de leur année d'entrée alors qu'il n'y en a que trois qui soient dans ce cas-là pour la seconde période : Oslo, Varsovie et Pittsburgh. Les années les plus dynamiques en termes de production sont celles qui se trouvent à la charnière entre les deux périodes, ce qui explique en partie les résultats obtenus. 74 % des agglomérations entrées avant 1999, republient au moins une fois avant 1999 (elles sont 80 % à republier au moins une fois avant 2007) quand seulement 27 % des agglomérations entrées après 2000 republient au moins une fois avant 2007.

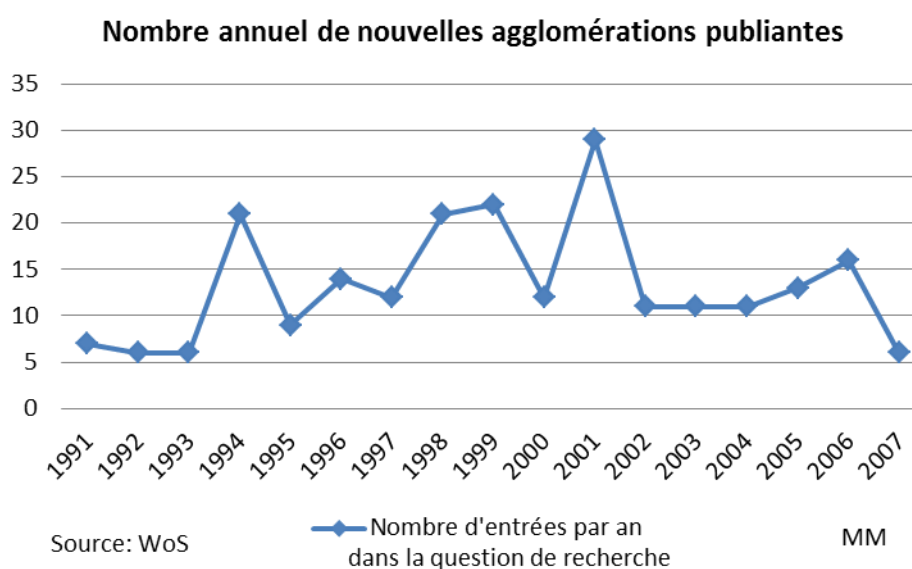


Figure 86 — Nombre annuel de nouvelles agglomérations publiantes sur la période étudiée.

Source : Web of Science

Cette évolution peut s'envisager comme le reflet de l'épuisement progressif de la question de recherche. Les nombreux entrants de la seconde période (Figure 85 et Figure 86) sont des « adoptants tardifs » qui, en majorité, ne s'investissent pas véritablement dans le sujet. Les cartes d'évolution des stocks de publications permettent également d'en rendre compte (Figure 87). On y voit aussi que les agglomérations entrées tardivement apparaissent dans l'entourage géographique des agglomérations pionnières, surtout les plus publiantes — qui ont fait paraître au moins cinq publications au cours de la seconde période. Trois agglomérations font exception : Cuernavaca (Mexique) et Oslo, dont l'activité se développe remarquablement, alors qu'elles appartiennent à des pays peu impliqués auparavant.

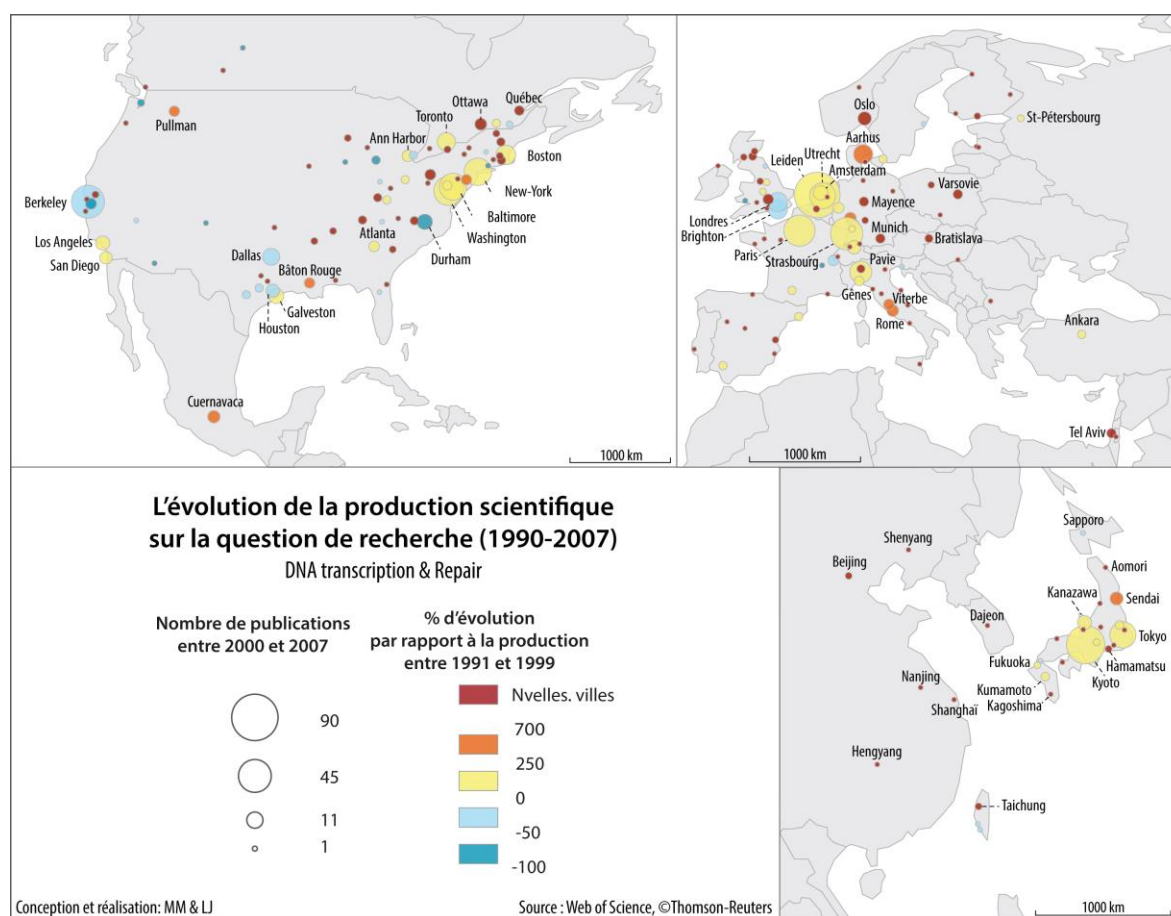


Figure 87 – Carte de l'évolution de la production scientifique sur la question de recherche (1990-2007). Source : Web of Science

D'après les informations recueillies auprès des protagonistes de cette question, non seulement il n'est pas facile de s'approprier les méthodes (notamment de clonage des gènes) et le matériel (y compris les cellules, mais aussi les modèles de souris mutantes) nécessaires à l'investigation de cette question de recherche mais en plus, la question tend à s'épuiser à partir du moment où tous les gènes impliqués dans le lien entre réparation et transcription de l'ADN ont été clonés. Ce sont les éléments d'explication, suggérés par les données et confirmés, lors de nos entretiens, par plusieurs biologistes travaillant sur cette question qui permettent de comprendre la timidité du processus de diffusion.

1.2. L'organisation spatiale de la question de recherche

Cette faible expansion du volume d'activité consacré à la question de recherche se comprend encore mieux lorsqu'on regarde les données de collaborations. À moins de collaborer avec les initiateurs de la question de recherche, il semble difficile de s'y investir

réellement. Cela nous ramène à la question de la circulation ou de l'échange de ressources, évoquée dans le chapitre précédent (Chapitre 10).

Ici, 45 % des publications ont été co-signées par au moins deux agglomérations différentes : 147 (41 %) dans la première période et 263 (47 %) dans la seconde. Alors que la procédure d'extraction des données ne le laissait pas prévoir, le réseau de co-publications obtenu à partir du corpus compte très peu de composantes non connectées à la composante principale : une seule dans la première période et quatre dans la seconde période qui correspondent à des collaborations très locales. Aussi, sur les 230 agglomérations considérées, il n'y en a que 25 qui sont isolées (6 dans la première période et 20 dans la seconde sont des agglomérations d'où n'a jamais été publié un seul article en collaboration avec une autre agglomération).

C'est que cette question de recherche, pour être investie, suppose de se mettre en relation avec des centres de recherche disposant de cellules spécifiques, à savoir les cellules de patients porteurs des maladies orphelines étudiées. À l'échelle européenne, des bibliothèques de cellules de ce type sont entretenues à Rotterdam, Paris, Pavie et Brighton (information confirmée par nos interlocuteurs lors des entretiens). Comme le démontre Dominique Vinck à l'occasion de l'évaluation d'un programme européen de santé, les objets intermédiaires jouent un rôle important dans la structuration des réseaux de collaboration scientifique en biomédecine – comme ici, les bibliothèques de cellules (Vinck, 1999). La nécessité de se procurer les cellules et le savoir-faire propre à la résolution de la question du lien réparation-transcription implique que les différents centres travaillant cette question dans le monde soient mis en relation, d'où le petit nombre d'isolats que l'on trouve. Par ailleurs, la centralité de certaines équipes est d'autant plus compréhensible qu'elles disposent de bibliothèques de cellule dont les autres équipes ont besoin pour pouvoir travailler sur la question. Les isolats, de leur côté, ont un niveau de production très bas (un ou deux articles publiés sur toute la période). Il s'agit en quelque sorte d'exceptions puisque, par ailleurs, dans le cadre de cette question de recherche, moins l'on est productif sur le sujet, plus il y a de chance que nos productions soient des co-productions. 80 % des 98 agglomérations qui ne publient qu'un seul article sur toute la période l'ont publié en collaboration avec une autre agglomération. Cela traduit la moindre autonomie de ces équipes, qui n'ont ni le savoir-faire, ni le matériel à disposition pour publier seules.

On observe effectivement que ce sont les agglomérations les plus productives qui ont tendance à produire le plus de publications en solo (non co-signées avec une autre agglomération). 70 % des 500 publications de ce type ont été produites par les 28 agglomérations les plus productives (ayant signées plus de 10 publications au total). Il y a toujours quelques contre-exemples : Pavie, Aarhus, Amsterdam et Kanazawa, parmi les aggloméra-

tions les plus productives, qui ne publient jamais seules. Mais il existe des explications. Les articles recensés à Aarhus (Danemark) sont en fait produits sous le patronage d'A. Bohr (le petit fils de N. Bohr) qui possède une double affiliation : l'Université d'Aarhus et le *National Institute of Aging* de Baltimore. Les autres cas sont moins évidents à justifier, seulement ils concernent des agglomérations qui sont intégrées dans des structures de collaboration denses : Pavie entretient avec Brighton une relation privilégiée sur ce sujet (elles sont impliquées dans des programmes de recherche communs), Amsterdam est proche des initiateurs néerlandais de la question de recherche et n'est donc pas très autonome vis-à-vis d'eux (Rotterdam-Leyde) et, de la même manière, Kanazawa s'inscrit dans le cluster japonais.

Tout comme ce sont les agglomérations les plus productives qui tendent à publier de façon autonome, il n'y a que les pays les plus productifs qui co-publient de façon autonome : autrement dit, les publications co-écrites par des agglomérations du même pays sont uniquement le fait des pays les plus publiants du corpus. Au cours de la première période, les huit pays à être dans ce cas sont les huit premiers producteurs de la période : USA, Pays-Bas, France, Japon, Royaume-Uni, Canada, Italie et Suisse. En plus des précédents, il y a cinq nouveaux pays « autonomes » dans la seconde période : l'Allemagne, le Brésil, l'Espagne, la Chine et la Finlande. On trouve, cependant, quatre pays produisant tout autant sur la période, sans pour autant entretenir de collaborations intra-nationales : le Danemark dont le bon niveau de publication est à rattacher au lien entretenu avec Baltimore ; la Norvège et la Pologne dont le bon niveau provient de l'activité intensive d'agglomérations « solitaires » (dans leur pays), et entrées tardivement sur le sujet (Oslo et Varsovie déjà évoquées) ; puis enfin, le Mexique. Les productions mexicaines sont toutes signées à Cuernavaca dont l'activité démarre en 1999 et reste régulière et autonome jusqu'en 2007, où l'on note une collaboration avec l'équipe de la transcription de Strasbourg.

L'évolution de la structure du réseau de collaboration semble caractéristique du passage d'une période d'émulation scientifique – d'innovation – à une période de diffusion – d'exploitation – des résultats de la question de recherche. En particulier, on constate qu'à l'exception des agglomérations américaines, les agglomérations du collectif collaboreront moins avec l'étranger dans la seconde période qu'elles ne le faisaient au cours de la première période.

Si l'on considère l'ensemble des liens, la part des liens intra-nationaux diminue dans le temps en nombre et en valeur. D'une période à une autre, le réseau se densifie et s'agrandit. Pris dans sa globalité, il paraît s'internationaliser. Mais, cette internationalisation apparente est uniquement le fait de l'ouverture des États-Unis à l'international au cours de la seconde période, ce qui concorde avec les tendances globales mesurées précé-

demment toutes disciplines confondues (Chapitre 9). En effet, en valeur des liens et en nombre de liens, la part des liens « intra-nationaux » (hors USA) sur l'ensemble des liens du réseau (hors USA) progresse d'une période à une autre (de 30 à 34 % en nombre et de 33 % à 48 % en valeur pour les liens supérieurs ou égaux à deux). Comme les États-Unis sont omniprésents dans le réseau de co-publications (les liens entre deux entités dont l'une au moins est américaine représentent en nombre 64 % de l'ensemble des liens de la seconde période et 56 % en valeur), et que les agglomérations américaines ont tendance à s'ouvrir à l'international entre les deux périodes, cela a pour conséquence d'occulter le phénomène de « re-nationalisation » intervenu dans la plupart des autres pays.

Ainsi, d'une période à l'autre, la part des collaborations intra-nationales augmente pour tous les pays « autonomes » de la première période à l'exception des États-Unis. Aux États-Unis, cette part passe de 53 % (la plus élevée avant la part japonaise de 41 %) à 35 % (devenant inférieure à la part japonaise, qui est passée à 47 %). Au niveau des agglomérations américaines, la part des co-publications intra-nationales progresse très faiblement pour Washington et New-York et chute considérablement pour la baie de San Francisco qui s'ouvre davantage à l'international (seulement un quart de ses co-publications était internationales au cours de la première période). Dans les autres pays, l'augmentation des liens intra-nationaux entre les deux périodes étudiée est sensible pour le Japon, les Pays-Bas, le Royaume-Uni et l'Italie. Cette évolution ne s'explique pas seulement par l'entrée de nouvelles équipes sur le sujet à proximité des pionnières (suivant le principe de contagion), mais s'explique aussi par la diminution en valeur de liens internationaux autrefois structurants pour la question de recherche.

Dans le réseau de collaboration, comme on pouvait s'y attendre vu la nature de l'objet étudié, les liens sont particulièrement instables ou éphémères. Si 87 agglomérations sur 119 présentes dans la première période publient de nouveau dans la seconde période, seulement 77 des 271 liens de co-publication de la première période se maintiennent dans la seconde période. Ces liens sont les liens structurants du réseau de co-publication et sont observables sur les graphiques où n'apparaissent des connexions entre agglomérations qu'à partir du moment où au moins deux publications ont été écrites conjointement sachant que le calcul des liens n'a pas été fractionné (Figure 88 et Figure 89). Sur ces représentations (Figure 88 et Figure 89), seuls les liens supérieurs à deux sont visibles ; les groupes de couleurs (sauf turquoise) correspondent à des sous-groupes connexes cohésifs ou *Islands* ; la surface des nœuds dépend de l'importance de leur indice d'intermédiarité ; et enfin, les agglomérations dont le label est précédé du signe « # » sont le produit du regroupement de plusieurs villes publiantes en agglomération.

Entre les deux périodes, 83 % de l'évolution de la valeur des liens s'explique par la progression de la valeur des liens « intra-nationaux ». En particulier, les équipes de Rot-

terdam et Leyde tendent à moins travailler avec l'équipe de Strasbourg, ce qui s'explique, justement, par l'épuisement progressif de la question de recherche dont ce trio était à l'origine. La collaboration privilégiée entre ces trois équipes était justifiée par la question de recherche puisqu'au départ, la question requiert la compétence de spécialistes de la transcription de l'ADN. Or, une fois le problème initial réglé et qu'il ne reste à traiter que des sous-questions dérivées de ce problème, l'équipe de Rotterdam n'éprouve plus le besoin d'activer ce lien. Cette hypothèse, d'abord inspirée par l'analyse de réseaux, a ensuite été validée par les spécialistes néerlandais du domaine.

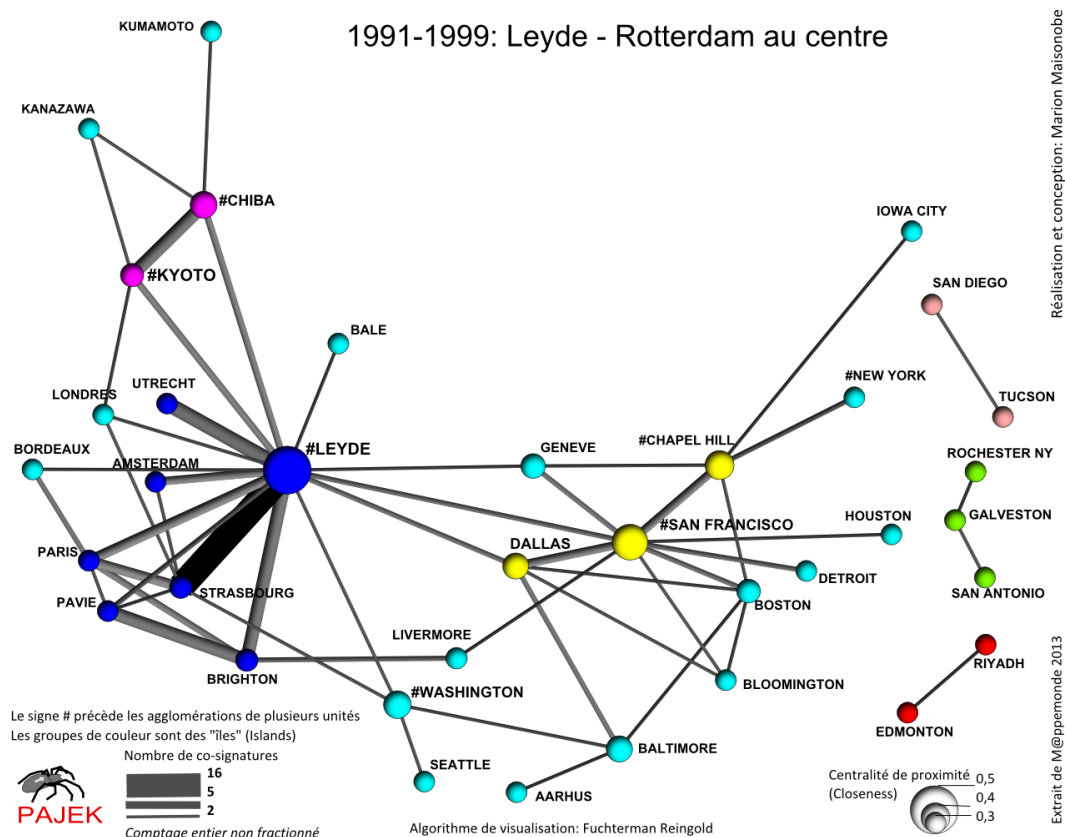


Figure 88 – L'évolution de la structure du réseau de co-publications dans le temps : 1991-1999 : Leyde-Rotterdam au centre (liens supérieurs à 2). Source : Web of Science

Le réseau de la première période montre une structure dont la hiérarchie est lisible. Il y a un groupe très cohésif autour des initiateurs de la question de recherche (le groupe européen) auquel se connecte d'un côté, le groupe japonais par l'intermédiaire d'Osaka et de l'autre, les Américains, qui constituent un groupe moins cohésif (les trois relais principaux sont Washington, la baie de San Francisco et Durham (#Chapel-Hill en Caroline du Nord qui fait le lien entre l'agglomération de New-York et le réseau européen *via* Genève).

Le réseau de la seconde période est plus complexe : il y a davantage de liens entre les membres des différents groupes : la structure d'ensemble est moins hiérarchisée. Paris, New-York et Pavie deviennent aussi centrales que l'étaient Leyde et Washington à l'origine (avec un indice de centralité de proximité¹³⁸ supérieur à 0,5). Le groupe néerlandais tend à se refermer sur lui-même : il forme une « île » à part c'est-à-dire, un sous-groupe connexe dont les nœuds sont davantage connectés entre eux qu'avec leurs voisins¹³⁹ (voir plus de précisions sur ce type de structure, dans la Partie 2. Chapitre 8.2.2).

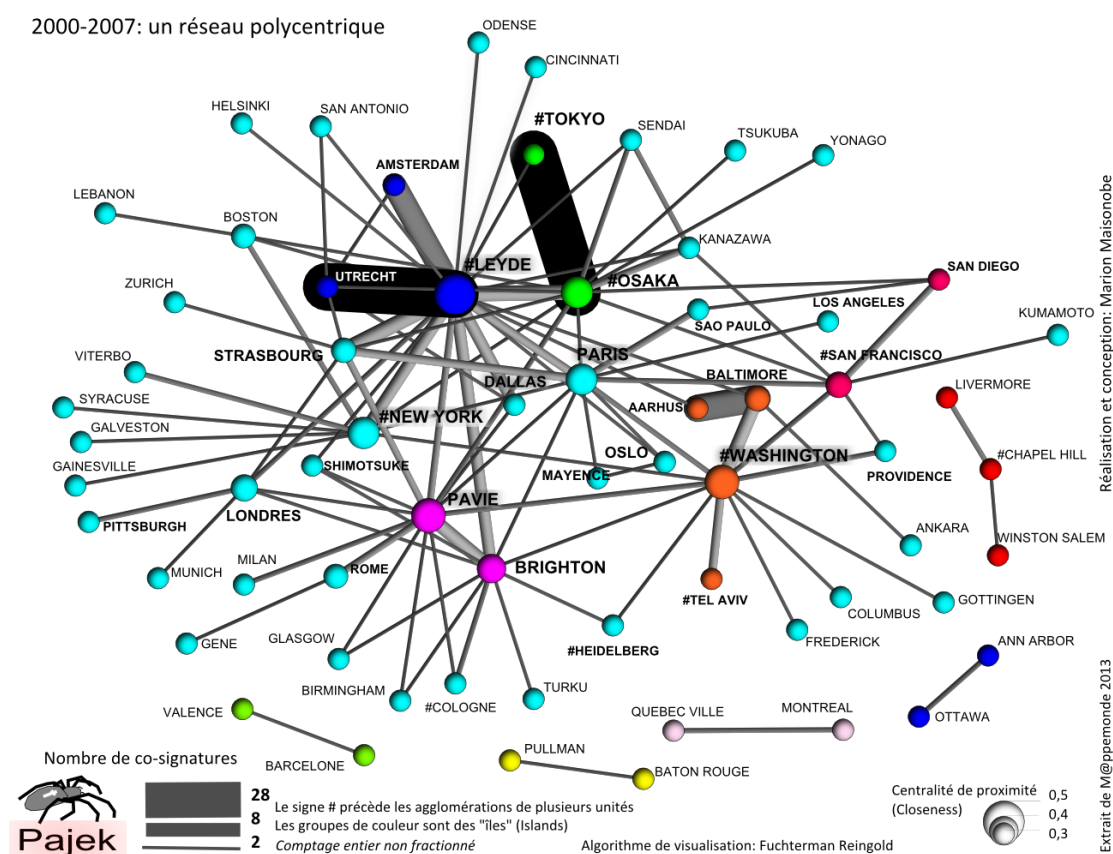


Figure 89 – L'évolution de la structure du réseau de co-publications dans le temps : 2000-2007 : Une organisation polycentrique (liens supérieurs à 2). Source : Web of Science

¹³⁸ La centralité de proximité ou « closeness centrality » : il s'agit de l'inverse de l'indice de Shimbel. Il se calcule pour un nœud donné i à partir de la distance de ce nœud i à tous les autres nœuds du graphe.

¹³⁹ La valeur de lien minimale dans l'île est plus grande que la valeur de lien maximale au voisinage de l'île.

Dans la seconde période, tout se passe comme si les équipes en tête des pays « autonomes » avaient contribué, à l'échelle nationale, à la diffusion de la question de recherche par contagion. Simultanément, ces équipes interagissent moins à l'international. C'est que l'intérêt stratégique de certaines collaborations tend à diminuer à mesure que s'épuise le potentiel d'innovation de la question de recherche. L'étude de cas d'Atkinson *et al.* sur le clonage d'un gène réalisé par plusieurs groupes de recherche en 1991 décrit un phénomène comparable. Elle montre, en se penchant sur le cas d'une équipe britannique, comment des relations mises en place par nécessité entre les membres d'un collectif tendent à s'émousser lorsqu'elles ne sont plus nécessaires pour la résolution du problème (Atkinson, Batchelor, & Parsons, 1998). L'idée d'Atkinson *et al.* est qu'au stade de partage initial succède la compétition, le « chacun pour soi ». Cette étude plus qualitative portant sur une équipe en particulier et ses partenariats est éclairante pour le cas qui nous intéresse (même s'il s'agit d'un domaine différent de la biomédecine).

L'analyse de réseaux permet d'observer que la dynamique de recherche fondée, pendant la première période, sur l'entretien de relations internationales – aborder la question de recherche suppose de se connecter à ses initiateurs (l'équipe de Rotterdam en particulier) suivant une logique « d'attachement préférentiel », c'est-à-dire une logique d'attraction pour les équipes les plus centrales – s'essouffle dans la seconde période au profit de logiques relationnelles relevant d'autres ressorts telles que l'« homophilie », l'entretien de relations avec des semblables et des proches en qui l'on peut avoir confiance. En l'espèce, cela se traduit par une intensification des liens intra-nationaux. Le renforcement de ces derniers bénéficie aussi d'une densification du réseau selon le principe de « *triadic closure* » : le fait de développer des relations avec les proches de ses proches (W. Powell *et al.*, 2005 ; Ter Wal & Boschma, 2009). L'analyse rejoint en ce sens celle de Leydesdorff et Rafols qui, comparant l'évolution des réseaux de co-publications pour deux technologies émergentes, montrent que, pour les deux technologies, la logique d'attachement préférentiel tend progressivement à diminuer au profit d'une logique de « petit-monde » (Leydesdorff & Rafols, 2011). Ce changement de stratégie est compréhensible si l'on tient compte du passage d'un stade de recherche à un autre. L'enjeu scientifique n'est plus le même dans la seconde période puisqu'à partir du moment où tous les gènes ont été clonés, il est davantage question de diffusion ou de ré-exploitation des résultats que d'innovation.

En somme, les résultats de cette analyse suggèrent que le passage d'un stade de recherche à un autre dans le cadre de l'investigation d'une question scientifique a une influence sur la diffusion spatiale de la question ainsi que sur l'organisation spatiale du réseau scientifique qu'elle dessine. L'observation de la production montre que la diffusion de la question s'organise essentiellement au voisinage géographique des pionniers.

Et s'il y a de nouveaux pays qui participent au cours de la seconde période, leur niveau de participation et son intensité ne témoignent pas en général d'un réel engagement mais plutôt d'une prise de connaissance de la question scientifique et éventuellement d'une exploitation des résultats auxquels elle a donné lieu. Ainsi, du fait de l'épuisement progressif de cette question, la diffusion reste relativement limitée dans l'espace géographique. Cette diffusion est d'autant plus contrainte que le sujet n'est pas abordable pour le premier venu. Il y a un coût à l'entrée qui ne se limite pas à acquérir le savoir-faire, mais suppose aussi de se procurer le matériel et les données nécessaires. Ce dernier réquisit suppose, sinon de disposer de cellules, au moins d'être en contact avec une équipe pourvue d'un stock de cellules. Pour cette raison, les productions « autonomes » restent essentiellement le luxe des contributeurs qui disposent sur place de tout ce dont ils ont besoin pour participer à l'avancement des recherches. Pour cette raison, le réseau va en se densifiant au profit de la centralité des zones disposant de ressources ; et rares sont les agglomérations qui n'y sont pas connectées à un moment donné.

Si des travaux antérieurs ont établi que la structure d'un réseau scientifique est sensible au passage d'un stade de recherche à un autre, ce cas d'étude permet d'explorer plus précisément comment évolue la structure spatiale d'un tel réseau. Il va dans le sens des travaux précédents qui montrent qu'une fois la question devenue moins brûlante, les collaborations stratégiques ou opportunistes tendent à diminuer au profit de relations de complicité ; ou plus exactement, il suggère que ce processus connu se traduit en termes géographiques par une préférence progressive pour l'entretien de liens domestiques, et donc pour le renforcement de la densité scientifique des systèmes nationaux. Ainsi, à mesure que la question s'épuise, en Europe comme au Japon, les agglomérations se replient sur leur contexte national et l'intensité de leurs relations internationales décroît considérablement. À rebours de tous les autres pays, les États-Unis ont tendance à s'ouvrir ce qui donne l'apparence, de prime abord, qu'un mouvement général d'internationalisation affecte le réseau. Dès le départ, les différents groupes américains (Washington-Baltimore ; la côte californienne ; le groupe texan ; l'agglomération new-yorkaise) sont peu connectés pour travailler sur cette question, ce qui suggère l'existence d'une compétition à l'intérieur des frontières américaines pour le partage des données et des résultats. Dans la seconde période, on observe une ouverture des agglomérations américaines à l'international, qui ne s'accompagne pas d'une intensification de leurs relations intranationales.

L'évolution du caractère structurant du cadre national selon le stade de recherche montre qu'il est essentiel de le mobiliser pour comprendre l'organisation des collectifs scientifiques, et qu'il ne devrait pas être éliminé des réflexions portant sur les réseaux scientifiques (Arvanitis, 2011).

Néanmoins, il faut se garder de surinterpréter le fait que les collaborations intra-nationales sont petit à petit privilégiées au détriment de certains échanges internationaux ; c'est-à-dire, qu'il faut s'empêcher d'en déduire que le principe d'homophilie se limite forcément au partage d'une nationalité, alors qu'il s'agit d'une logique d'attachement qui favorise l'entretien de relations entre individus ou groupes partageant une qualité commune quelle qu'elle soit. Ainsi, l'évolution observée indique tout au plus que la présence d'une équipe porteuse de la question de recherche dans un pays donné a favorisé le développement de travaux en relation avec cette question au sein d'autres équipes du même pays. Ce phénomène peut s'expliquer par le fait que certains systèmes nationaux sont organisés de telle sorte que nombre de carrières de chercheurs et d'échanges s'organisent à ce niveau institutionnel. Mais, il existe aussi des mécanismes institutionnels structurants à d'autres niveaux, qui pour leur part favorisent l'entretien de relations internationales : la double affiliation d'A. Bohr en témoigne, ainsi que l'organisation soutenue par des programmes et des bourses de recherche d'un réseau de collaboration européen. Aussi, on va le voir, le fait que beaucoup de post-doctorats soient effectués à l'étranger en biomédecine a facilité les échanges internationaux sur cette question de recherche et a pu contribuer à renforcer la cohésion du domaine de la réparation de l'ADN à différents niveaux.

Au lieu de se limiter au cas particulier d'une équipe et de son contexte relationnel, comme dans l'étude d'Atkinson *et al.* (*op. cit.*), notre approche permet de saisir l'évolution de toutes les équipes en même temps (au prix évidemment d'une perte de finesse qu'il faut accepter pour saisir la logique d'ensemble), et de repérer les systèmes de villes qui comme aux États-Unis, vont à contre-courant des autres ensembles urbains. Aussi, en se limitant à une question de recherche bien définie et circonscrite, on peut beaucoup mieux tenir compte des « objets intermédiaires » pour interpréter la géographie initiale et son influence sur les évolutions observées. Au bout du compte, cette étude montre qu'il y a un réel intérêt à considérer l'espace géographique, sinon pour le décrire ou l'expliquer, au moins en tant que variable pour saisir l'évolution des collectifs scientifiques dans le temps. Elle conforte enfin l'idée que les groupes scientifiques et l'entretien de relations ne sont pas stables dans le temps et gagnent à être étudiés en dynamique.

2. Du point de vue d'une chercheuse toulousaine

« Un cercle constitué en fonction des points de vue objectifs peut se construire par-dessus un cercle organiquement immédiat : la cohésion primitive du groupe familial est modifiée lorsque l'individualité d'un de ses membres l'intègre à des cercles plus larges ; l'un des exemples les plus élevés est la "république des savants", cette association est à demi-idéale, à demi-réelle, en vue d'une fin aussi universelle que la connaissance en général, de toutes les personnalités, qui par ailleurs appartiennent aux groupes les plus divers quant à la nationalité, les intérêts personnels et généraux, la position sociale etc. » (Simmel, 1908b, p. 413)

2.1. « Faire partie de la légende », puis revenir en France

Que ce soit pour définir les corpus bibliographiques, pour interroger la pertinence des résultats obtenus et pour valider les interprétations, nous avons sollicité à plusieurs reprises une informatrice privilégiée (AM), elle-même investie dans la question de recherche sur le lien entre réparation et transcription de l'ADN. Celle-ci a manifesté dès le départ un grand intérêt pour les études que nous voulions mener sur l'expansion géographique de son domaine de recherche. Son sentiment à l'égard de ce phénomène était celui d'une complexification récente, d'une remise en cause des anciennes hégémonies au profit de nouveaux pôles d'activité comme Singapour, brouillant peu à peu la représentation qu'elle se faisait de l'organisation de sa propre communauté.

La comparaison de la carte des recherches sur la réparation de l'ADN entre le stade d'émergence et aujourd'hui a montré que cette impression était justifiée. En particulier, la production chinoise a pris une place inattendue dans le milieu des recherches sur la réparation de l'ADN en général. Ceci-dit, la question de recherche investie par AM au début des années 2000 apparaît peu concernée par cette évolution. En effet, cette question est dominée dès l'origine par ses initiateurs à Rotterdam et Leyde et bien qu'ils en viennent à partager avec d'autres équipes une position centrale sur cette question dans les années 2000, ces dernières localisées à Washington, Paris, Osaka, Brighton et Pavie sont impliquées depuis longtemps dans le domaine de la réparation de l'ADN. En particulier, les équipes de Brighton et Pavie collaboraient déjà régulièrement avec les chercheurs de Rotterdam et Leyde grâce, à l'origine, aux financements EURATOM. L'évolution de la

position des équipes de Rotterdam et Leyde dans le réseau de collaboration sur la question de recherche est donc le reflet des échanges que les membres de ces équipes ont activé avec leur cercle traditionnel, mais aussi, on va le voir ici plus précisément, de l'attractivité provoquée à l'échelle européenne par les retombées de la découverte de 1993. Cependant, l'implication de nouveaux compétiteurs, qui auraient pu chercher à venir bousculer cette organisation bien établie, ne ressort pas de nos observations. Au contraire, les nouvelles équipes qui apparaissent à cette occasion émanent des espaces centraux de la spécialité et restent dans leur sillage.

La carrière d'AM est révélatrice à ce titre. Après avoir étudié la géographie de sa question de recherche, et plus largement de sa spécialité de recherche, se pencher sur la trajectoire d'AM et son « univers de référence » permet d'envisager le lien entre son positionnement cognitif et géographique. Au-delà, son parcours aide à comprendre, au plus près de la pratique de recherche, comment de nouvelles connexions peuvent s'établir entre lieux et comment de nouvelles équipes peuvent se former, participant à la densification des réseaux de collaborations scientifiques, généralement observée à des niveaux plus agrégés. Dans cette perspective et suite aux nombreux échanges que nous avons eu sur une base plus informelle, et l'aide qu'elle nous avait apportée pour rencontrer une partie de ses collègues néerlandais, cette dernière nous a accordé un entretien d'une durée d'environ 2h30 dans son laboratoire, à Toulouse, le 10 octobre 2012.

L'entretien s'est déroulé en reprenant et en adaptant la grille définie et décrite par Béatrice Milard dans son habilitation à diriger des recherches (Milard, 2011a), dont l'exploitation a abouti sur d'importants résultats en sociologie (Milard, 2010, 2011b, 2014). En particulier, interroger les chercheurs sur leur activité de production scientifique a montré que « les publications scientifiques sont l'occasion d'entretenir et de mettre en place des relations et des collectifs, qu'elles contribuent à organiser et à pérenniser. ». Certaines de ces relations sont activées au fil des parcours de recherche, d'autres sont des relations latentes caractéristiques de la position que les auteurs occupent ou souhaiteraient occuper dans l'espace scientifique, d'autres encore ne font que traduire l'appartenance à des communautés de pratiques (l'utilisation de telle méthode, de telle technique etc.). Se pencher sur la nature des relations qui sont ainsi mobilisées ou provoquées par la médiation d'une publication, en distinguant parmi les différents types de liens sociaux sous-jacents, est un moyen de voir comment la publication révèle autant qu'elle donne vie à un assemblage de personnes, d'idées, de pratiques, de lieux original qui fait sens du point de vue du chercheur interrogé. Or, comme il s'appuie partiellement sur des liens sociaux préexistants, cet assemblage ne peut se comprendre qu'au regard de la trajectoire du chercheur publiant, et notamment de sa trajectoire géographique et de la représentation géographique (ou carte mentale) qu'il se fait de sa spécialité, qui sont ici au

cœur de notre réflexion. Aussi, la méthode d'entretien a donc été un peu adaptée en fonction de l'objectif recherché, à savoir : situer la trajectoire et la position d'AM dans l'espace géographique de la question de recherche précédemment étudiée, et au-delà dans l'espace géographique de la communauté de la réparation de l'ADN ; et dans un second temps saisir la manifestation de ce positionnement et son rôle dans la configuration d'un assemblage original à travers l'exemple d'une publication d'AM, sélectionnée par ses soins.

Le parcours d'AM a commencé en région parisienne où elle a réalisé sa thèse à la fin des années 1990. Il s'agissait d'étudier les mutations d'une protéine dans les tumeurs de patients atteints de *Xeroderma Pigmentosum* (XP). La thèse relevait ainsi du domaine de la mutagenèse, l'étude des mutations, et pas véritablement de la réparation de l'ADN. Le travail consistait à observer l'évolution des tumeurs, mais pas à comprendre les mécanismes en jeu derrière ces évolutions. Après avoir soigneusement regardé et comparé une centaine de tumeurs, AM est arrivée à en tirer des résultats intéressants qui ont été valorisés dans *Cancer Research*, une revue à fort impact. Mais cet aspect des recherches sur les mutations l'intéressait beaucoup moins que l'approche « mécanistique » qu'elle connaissait bien à travers ses lectures : « J'avais envie de comprendre ce qu'il y a dans le mécanisme... Et la thèse avait été une épreuve, je me demandais si j'étais vraiment faite pour ça. ». En lisant les articles de l'équipe de Rotterdam sur les mécanismes de réparation de l'ADN, AM décréta que s'il fallait continuer ce serait là-bas ou nulle part : « c'était des mythes pour moi. » ; « j'avais envie de travailler sur la mécanistique et faire partie du mythe. ». Les échanges avec l'étranger furent limités pendant la thèse, et il n'y eut par exemple que des français dans son jury de thèse, mais dès qu'elle s'est « mis en tête de trouver un post-doc en Hollande », AM est entrée en contact avec le chef d'équipe de Rotterdam, et s'est rendue à un congrès pour le rencontrer.

Une fois le post-doc décroché, elle y est restée 7 ans parce que « ça marchait bien ». Ça a été l'occasion de définir son sujet : « créer une souris OGM et analyser les résultats ». Cette fois encore, elle est parvenue à publier un article dans une revue prestigieuse : *Nature*. Son article de 2004 reste une satisfaction : « C'est l'article qui m'a faite rentrer dans la légende, mais ce n'est pas celui dont je suis le plus fière parce qu'on savait ce qu'on devait faire. ». Ensuite, ne voyant pas d'évolution possible en termes de carrière, elle en est arrivée à se dire qu'il faudrait « changer d'endroit ou de boulot ». À ce moment-là, une, puis deux opportunités en France se sont présentées et il a fallu faire le choix entre Paris et Toulouse, ce qui n'était pas facile. En 2008, encore à Rotterdam, elle passe le concours CNRS et obtient une « *grant* en ATIP », un programme aidant les jeunes chercheurs à mettre en place leur équipe. C'est alors que pour obtenir rapidement des résultats, elle se voit contrainte de faire un peu évoluer son sujet : « À Rotterdam, je tra-

vaillais pendant 7 ans sur une souris précise, mais travailler sur les souris, c'est trop long et pour commencer une équipe, il faut des résultats rapides », sans compter que les moyens à disposition n'étaient pas les mêmes à Toulouse pour « faire du *in vivo* ». Bien que « toujours dans la mécanistique », la nouvelle approche qu'elle privilégie consiste à « chercher des réponses dans les cellules de la souris (méthode *proteomics*) *in vitro*, *in cellulo* », plutôt qu'*in vivo* comme elle le faisait avant.

Néanmoins, le travail antérieur n'est pas abandonné, et il fait l'objet d'une publication signée entre Toulouse et l'ancien laboratoire d'AM dès la première année de son arrivée. C'est cette publication qui est, en 2012, celle dont elle est la plus fière et qu'elle choisit comme publication à présenter et commenter en cours d'entretien. Elle considère cet article de 2009 comme « le résultat d'un *challenge*, le résultat de 7 ans de souris... le morceau d'ADN à mettre dans la souris, j'ai mis un an pour le faire ! Il fallait réfléchir à la base d'ADN près (à l'alphabet près). ». Toutes les cellules de la souris portent la protéine XPB taguée avec YFP pour « *yellow fluorescent protein* » qui est une méthode assez nouvelle permettant d'observer la dynamique. Le choix de cette méthode a été réfléchi, décidé après mûre réflexion parce qu'il y aurait des conséquences à long terme et que les résultats seraient longs à obtenir. Dans un tel cas de figure, il faut évidemment bien choisir son protocole. Le « résultat de 7 ans de post-doc » est publié dans cet article. Par ailleurs, elle nous apprend qu'un autre article a été dérivé de celui-ci mais n'a, pour sa part, pas encore été soumis en 2012. Il porte sur la quantité de protéines par cellule et envisage ces données comme marqueur pour étudier le facteur TFIIH de transcription et réparation de l'ADN. D'après AM : « il sera publié, c'est juste qu'il y a d'autres priorités... d'autant qu'on ne peut pas être dépassé (doublé), parce qu'on est les seuls à avoir cette souris et donc à pouvoir faire ce type de mesures sur elle. ».

Ce travail « très précurseur », publié en 2009, était encore « très peu cité » en 2012. D'après AM, c'est normal qu'il faille du temps pour un tel article. En Juin 2015, il a reçu 29 citations, a été cité en moyenne 4 fois par an depuis sa parution et figure à la dixième place parmi les publications les plus citées d'AM (Source : WoS). L'impact « limité » de l'article (relativement à son importance) avait été anticipé. L'article avait été proposé à *Nature* et *Science* mais n'avait pas pu y être publié justement pour cette raison : « L'éditeur de *Science* nous l'a demandé mais finalement ça ne pouvait pas passer parce que ce serait peu cité dans les prochaines années. ». En fait, les résultats de l'article s'appuient sur un « outil que personne n'a », mais d'autres équipes travaillent à le reproduire : « on nous a demandé les souris, les cellules » ; « un article de 2012 est sorti qui fait la même chose, ils nous avaient contacté bien avant, pendant qu'on était encore en train de faire notre expérience. ». Il faut ajouter enfin, avant d'en venir à l'analyse du contexte de collaboration de cette publication, puis de l'ancrage spatial de cette publication dans un ré-

seau de références et dans un univers de référence, que, selon elle, ce travail était de toute façon plutôt susceptible d'avoir de l'impact « dans le milieu de la transcription. ».

2.2. Une publication co-écrite entre Rotterdam et Toulouse

En biologie moléculaire, lorsque le nom de l'auteur d'une publication est suivi d'une étoile : « c'est pareil qu'en chimie. », ce qui veut dire que le premier qui a l'étoile est celui « qui a fait le travail », et le dernier, « celui qui a mis l'argent et les idées, le directeur. ». Avoir l'étoile, implique d'être auteur référent (« *corresponding author* »), c'est-à-dire l'auteur auquel s'adresser « pour demander le plasmide ou la souris ». Au milieu de la liste des auteurs figurent les contributeurs secondaires. Au total, l'article choisi par AM comporte 13 auteurs et 4 adresses : une à Rotterdam, deux à Toulouse (le laboratoire CNRS et l'Université), et une dernière à Leyde. Les explications apportées par AM en entretien nous font comprendre que ces liens de co-signatures sont sous-tendus par des mobilités entre ces lieux. AM a soumis l'article fin 2009 au moment où elle était encore post-doctorante senior à Rotterdam, puis s'est occupée des révisions de l'article depuis Toulouse : « l'article a été 9 mois en révision, on a fini l'expérience à Toulouse. ». Pour cette raison, AM a déclaré une affiliation multiple, ainsi que son compagnon et co-auteur qui a suivi la même trajectoire et s'est chargé avec elle d'analyser les résultats. Plusieurs autres auteurs étaient en transition entre deux postes au moment de la publication :

- la 4^{ème} auteure, à Bordeaux auparavant, est arrivée à Toulouse en même temps qu'AM, en tant que post-doctorante, et a participé aux travaux réalisées pendant les révisions ;

- la 5^{ème} auteure était en stage de Master avec AM à Rotterdam, puis a suivi AM à Toulouse pour sa thèse (soutenue en 2013), et se trouve désormais en post-doctorat à Rotterdam. Début 2009, cette dernière était encore aux Pays-Bas, si bien qu'AM a pu l'envoyer auprès d'une chercheuse en poste à Leyde pour répondre à une critique des referees (« arranger un aspect du papier », « apprendre une technique »). Aussi, la chercheuse qui lui a enseigné la technique (MF) a été « ajoutée dans les signataires entre les procédures de révision », devenant 11^{ème} auteure. Peu de temps après la publication, MF a quitté les Pays-Bas pour monter sa propre équipe en Grèce, son pays d'origine, transition importante dans sa carrière qui a été facilitée par la bourse européenne qu'elle a décroché peu de temps après, en octobre 2012 (bourse ERC).

- la 6^{ème} auteure était en post-doctorat en même temps qu'AM à Rotterdam. Elle a suivi quelque temps AM à Toulouse, avant de partir pour Barcelone en bio-informatique.

- la 8^{ème} auteure, enfin, était à Rotterdam au moment de la publication, mais venait de Paris, et y est retournée entre temps.

La moitié des auteurs, surtout des jeunes femmes, ont donc traversé une expérience de mobilité accompagnée d'un changement de statut à la période où l'article a été publié. Les autres auteurs, en contrepoint, avaient déjà un poste fixe au laboratoire de Rotterdam : deux ont participé à l'aspect technique de la publication (le 2^{ème} et le 7^{ème} auteur), deux autres ont contribué à l'expérience (9^{ème} et 10^{ème} auteur), et les deux derniers signataires (12^{ème} et 13^{ème} auteur) étaient respectivement le directeur de laboratoire et le directeur d'équipe. La publication marque un trait d'union entre les années de post-doctorat d'AM dans un grand laboratoire et son nouveau statut de chef d'équipe. Elle témoigne à travers le retour d'AM en France (et celui de MF en Grèce) d'un processus de diffusion de la thématique et des méthodes de travail néerlandaises au reste de l'Europe. L'apparition de ces nouvelles équipes en 2009 et 2010 découle directement du phénomène de rayonnement et d'attractivité provoqué dans les années 1990-2000 par le dynamisme des laboratoires de Rotterdam et Leyde en matière de réparation de l'ADN. Ce processus se traduit par la formation de jeunes chercheurs au stade du post-doctorat, qui ont ensuite réinvesti le sujet dans un autre pays en maintenant des liens avec les Pays-Bas. Aussi, les trajectoires des auteurs de la publication s'inscrivent toutes dans un espace européen et le travail sur lequel s'appuie la publication a bénéficié de deux financements européens provenant de l'*European Molecular Biology Organization* (EMBO) et de l'Union Européenne. L'Europe apparaît donc comme un niveau d'organisation structurant pour ce collectif de recherche — même si les financements dont ont bénéficié les auteurs de la publication sont également venus d'autres niveaux institutionnels : l'Université Erasmus de Rotterdam, l'Organisation Néerlandaise pour les Sciences Médicales (ZonMW), l'Association International de Recherche sur le Cancer (rebaptisé entre temps « *Worldwide Cancer Research* » mais dont le comité scientifique est essentiellement composé de chercheurs européens¹⁴⁰), et la Région Midi-Pyrénées. D'un point de vue institutionnel, cette publication est donc le résultat d'une combinaison de dispositifs néerlandais, français et européens.

Sachant qu'on trouve aussi des spécialistes de la question de recherche en Amérique du Nord et au Japon, il n'est par ailleurs pas étonnant qu'un Américain ait été chargé de l'édition scientifique de la publication (choisir les referees) et que quelques Japonais soient cités parmi les références ; mais aucun chercheur de ces deux espaces géographiques n'a participé à la conception et à la réalisation du travail, et donc n'a collaboré formellement avec les auteurs.

Bien que certains des auteurs de la publication n'aient peut-être pas directement travaillé ensemble, AM a pour sa part travaillé avec chacun d'entre eux et occupé le même espace géographique qu'eux à un moment donné (situation de co-présence). Dans les faits,

¹⁴⁰ Source : URL : <http://www.worldwidecancerresearch.org/about/scientific-advisors>

comme on l'a vu précédemment (Chapitre 10), le fait de co-écrire un article implique généralement d'avoir été en « proximité physique » à un moment donné mais cela ne se manifeste pas forcément par le fait de partager la même affiliation. À titre illustratif dans le cadre de cette publication, le passage de la future doctorante d'AM à Leyde pour apprendre une technique ne s'est pas manifesté par un changement d'affiliation de la jeune chercheuse, qui était alors en stage à Rotterdam, c'est-à-dire à moins d'une heure de Leyde en train. Des déplacements et des expériences de mobilités, pas forcément formalisés, ont ici permis de faire la jonction entre les trois localités de signature. Comme c'était aussi le cas pour les trajectoires transnationales repérées au stade d'émergence de la spécialité, ces déplacements ont été plus souvent le fait de jeunes chercheurs qui n'ont pas encore le statut de permanent dans un laboratoire. Il est important de le signaler, car les études sur la mobilité tendent parfois à insister sur la mobilité des chercheurs « stars ». En particulier, c'est sur ce type de mobilité que s'est penchée Grit Laudel dans son étude de spécialité fondée sur l'analyse des changements d'affiliation (Laudel, 2003 ; Laudel, 2005), alors que la dynamique d'une spécialité repose aussi en grande partie sur sa capacité à attirer de jeunes recrues et à les former.

En dehors des travaux de sciences de l'innovation qui insistent sur l'efficacité et l'avantage de la mobilité en termes de productivité (le « *brain gain* » envisagé par Saxenian, 2006 ; Agrawal *et al.*, 2011) et d'impact scientifique, notamment à partir des données de brevets (Jaffe, Trajtenberg, et Henderson, 1993 ; Almeida et Kogut, 1999 ; Breschi et Lissoni, 2009), il y a désormais tout un ensemble de travaux qui se préoccupent des déterminants de la mobilité (Agrawal, Cockburn et McHale, 2006 ; Bernela, Bouba-Olga et Ferru, 2013). Ces dernières recherches tendent à montrer l'importance des réseaux personnels (des liens sociaux) pour expliquer les trajectoires géographiques des chercheurs. À ce titre, ces travaux s'inscrivent dans la continuité des études entamées par des sociologues américains à la fin des années 1980 qui montraient le lien entre l'avancement de la carrière et la mobilité en tenant compte de l'influence des facteurs humains tels que l'âge ou le sexe (Allison et Long, 1987). Toutefois, comme on manque de données sur la mobilité des scientifiques, puisqu'elle est mal renseignée par les organismes de statistiques internationaux, la plupart des études sont limitées à une zone géographique donnée, s'appuient sur des échantillons de population qui ne permettent pas d'avoir un aperçu d'ensemble des flux de chercheurs à l'échelle mondiale et de leur dynamique (Cañibano, Otamendi, & Solís, 2011). C'est d'autant plus regrettable que les pratiques de mobilité peuvent varier en fonction du pays d'origine mais aussi de la discipline (*ibid.*), voire même de la question de recherche. En effet, comme le montre l'exemple que nous étudions ici, à la fin des années 1990, Rotterdam apparaissait, depuis la France, être une destination in-

contournable pour qui souhaitait étudier les mécanismes de réparation de l'ADN par excision de nucléotide.

Aussi, les parcours de mobilité peuvent dépendre de facteurs conjoncturels et les chercheurs suivent parfois, lorsqu'un pays est « en crise », la logique des flux migratoires du reste de leurs compatriotes (Meyer, Kaplan, & Charum, 2001 ; Barré, Hernández, Meyer, & Vinck, 2003). En particulier, comme nous l'avons détecté précédemment, certains spécialistes de la réparation de l'ADN avaient quitté l'URSS pour Israël dans les années 1970 ; et, d'autres, dans les années 1950-1960, avaient quitté l'Allemagne pour les États-Unis. Plus près de nous, la collègue d'AM a dû faire une partie de sa carrière en Angleterre et aux Pays-Bas pour pouvoir se faire une réputation qui lui permettrait de retourner dans son pays, la Grèce. En effet, c'est grâce à son passage dans les laboratoires les plus fameux de sa spécialité, qu'elle a pu obtenir une bourse européenne suffisamment importante pour créer une nouvelle équipe dans son pays d'origine. Sans cette bourse, elle admet qu'elle n'aurait pas pu en faire autant car les universités grecques manquent actuellement de fonds¹⁴¹. Quand les mobilités sont ainsi contraintes par l'incitation à faire un post-doctorat à l'étranger ou le manque de moyens et de postes dans un environnement institutionnel donné, l'aspiration possible au retour peut se manifester et être mesurée par le maintien des relations scientifiques avec le pays d'origine (Baruffaldi et Landoni, 2012). Il faut dire que la mobilité n'est pas toujours vécue positivement par les individus. Et de plus, contrairement aux idées reçues, elle n'est pas indispensable à la qualité d'une carrière scientifique. En effet, les exemples de carrières de chimistes étudiés par Milard ont montré qu'il est possible de réaliser des contributions scientifiques importantes, et même d'arriver à faire évoluer son sujet de recherche et son cercle de collaboration tout en restant dans le même laboratoire (l'exemple de Sophie, *op. cit.*). Dans ce dernier cas, les collaborations s'appuient davantage sur des mobilités temporaires (qu'on peut difficilement saisir à travers les données de publications), et sur la densité des liens avec l'environnement local. La mobilité n'est donc pas une sinécure dans l'exercice de la recherche mais reste un phénomène intéressant à étudier, surtout si l'on veut comprendre la géographie de certains domaines de recherche et plus généralement l'évolution de la géographie des activités scientifiques. Pour l'instant, elle est encore trop souvent étudiée en relation avec son efficacité, alors qu'elle gagnerait à être perçue comme un vecteur parmi d'autres (les correspondances, les circulations de matériels et spécimens...) pour la diffusion de certaines idées et techniques de recherche et comme mécanismes sous-jacents à la formation de collectifs scientifiques.

¹⁴¹ Entretien de Maria Fousteri, accessible en ligne. Source : URL : http://www.researchresearch.com/index.php?option=com_news&template=rr_2col&view=article&articleid=1346732

En plus d'être typique d'un phénomène d'expansion géographique de la question de recherche étudiée, l'article choisi par AM donne un exemple intéressant de collaboration internationale, qui présente, il faut le souligner, la particularité d'être exceptionnel dans la carrière de la chercheuse. En effet, elle n'avait pas changé de statut et de pays depuis 7 ans, n'en a toujours pas encore changé depuis ; et il se peut très bien, comme elle dispose désormais de sa propre équipe, qu'elle n'en change plus du tout¹⁴². Dans l'ensemble, le parcours d'AM et sa publication donnent à voir des situations qui rejoignent celles décrites par Milard à partir des nombreux entretiens qu'elle a mené avec les 32 chimistes d'un laboratoire toulousain réputé, et qui l'ont conduite à faire le diagnostic suivant : « Nous retrouvons bien des mobilités qui s'inscrivent dans des collaborations internationales dont elles découlent ou qu'elles initient, mais celles-ci ne concernent pas des chercheurs les plus reconnus, mais plutôt des jeunes chercheurs post-doctorants gérant comme ils peuvent ces périodes de précarité. Nous n'observons pas non plus de hiérarchie de prestige scientifique entre les relations locales, nationales et internationales. Nous voyons plutôt une toile de liens de toutes sortes, des complémentarités durables ou de circonstance, et une sorte de banalité relationnelle, quelle que soit la localisation des protagonistes, dans un domaine où le travail collectif est la règle depuis longtemps. » (Maiso-nobe *et al.*, à paraître).

Tandis que les collaborations activées au moment de la publication sont le fruit de la rencontre de plusieurs trajectoires, et du « déménagement » de l'expérience scientifique d'AM accompagnant son changement d'affiliation, la publication révèle aussi tout un ensemble de relations, latentes, à travers son réseau de références (l'ensemble des références qu'elle cite), et au-delà son univers de référence (l'ensemble des publications mobilisant les mêmes références qu'elle). C'est sur ces deux derniers corpus de publications, et plus particulièrement sur leur ancrage géographique, que nous allons fonder le dernier volet de notre analyse.

2.3. La géographie du réseau et de l'univers de référence de la publication

Hormis les collaborations qui sous-tendent la production de la publication, cette dernière donne à voir tout un ensemble de relations à travers les références qu'elle mobilise. En effet, les références citées participent au positionnement de la publication dans un collectif de recherche. Demander à la chercheuse responsable de la publication de qualifier

¹⁴² Dans les faits, l'équipe d'AM est, à l'été 2015, en pleine phase de déménagement vers un nouvel institut à l'Université Claude Bernard de Lyon. L'objectif de ce déplacement est pour AM et son équipe de reprendre des études *in vivo*.

oralement les relations qu'elle entretient avec chacune de ces références, suivant la méthode proposée par Béatrice Milard (2011a et 2011b, *op. cit.*), permet de voir comment se situe son travail, sa trajectoire et son espace de collaboration dans un ensemble sociocognitif plus vaste. En effet : « S'intéresser aux entourages citationnels permet de déborder le cadre organisationnel en tenant compte des relations moins formelles. Le spectre des relations est alors plus large et inclut différents types d'acteurs ou groupes impliqués dans la recherche : les collaborateurs, mais aussi les auteurs influents, les compétiteurs, les incontournables et ceux qui le sont moins, les contemporains, mais aussi des auteurs du passé... » (Milard, 2014).

Indépendamment des relations sociales et d'interconnaissances entretenues avec les auteurs de certaines des références citées, il importe aussi d'analyser et d'interpréter l'ensemble de références au regard de l'objectif de la publication et de sa place dans la carrière de la chercheuse. Ici, il s'agit d'une publication charnière pour AM, qui marque la fin et l'aboutissement de son post-doctorat à Rotterdam. Et la contribution scientifique de la publication a été d'appliquer une méthode existante pour marquer et suivre une protéine dans toutes les cellules d'un animal *in vivo* sur une souris mutante, conçue à Rotterdam, qui exprime la mutation XPB (pour *Xeroderma Pigmentosum* du groupe B). Cela a permis d'étudier la dynamique de la protéine et de mesurer des différences dans cette dynamique, en fonction des types de cellules où la protéine est intervenue. Dès lors, les qualifications apportées par AM sur chacune des références citées dans la publication montrent comment sa trajectoire personnelle et sa culture scientifique ont été mobilisées au service de cet objectif. Au-delà, puisque c'est ici le lien entre la trajectoire géographique d'AM et la géographie de son univers sociocognitif qui nous intéresse, les informations recueillies lors de l'entretien ont été mises en regard avec l'inscription spatiale des références de la publication.

En effet, l'ensemble des 63 références citées dans la publication a été représenté ci-dessous sous la forme d'un réseau : chaque nœud du réseau est une référence (le numéro correspond à l'ordre d'apparition de la référence dans le texte), et les références sont liées lorsqu'elles partagent au moins une localité en commun¹⁴³ (Figure 90). Une partie des commentaires de la chercheuse ont été reproduits sur la Figure en violet. Certaines relations ont été accentuées : en particulier celles reliant Rotterdam à Leyde et Rotterdam à Strasbourg, ce qui a permis de faire ressortir plusieurs sous-groupes matérialisés par des cercles : en jaune, le groupe des références signées exclusivement depuis Rotterdam ; en rouge, le groupe de références signés depuis Rotterdam et Leyde, éventuellement avec la collaboration d'autres équipes comme Brighton, Pavie et Strasbourg ; en vert, le groupe

¹⁴³ Dans le cadre des analyses présentées dans cette section, nous n'avons pas procédé à une agrégation des localités déclarées par les auteurs au niveau des agglomérations urbaines.

de références signées depuis Rotterdam et Strasbourg ; et en bleu, le cercle de références qui ont en commun d'avoir été signées depuis le Royaume-Uni, et parfois depuis les même localités (lorsqu'il y a un lien entre elles).

Réseau de références qualifié de l'article commenté en entretien

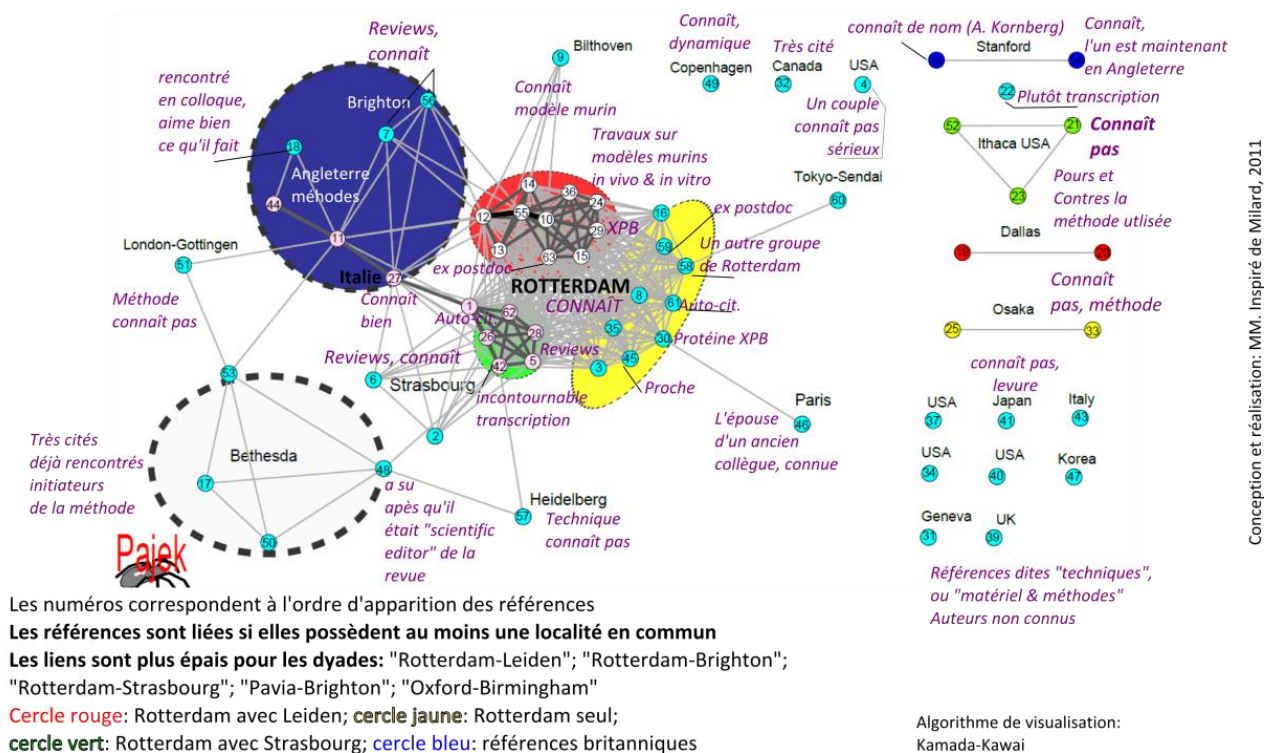


Figure 90 – Réseau de références qualifié de l'article scientifique d'AM commenté en entretien

41 références appartiennent à la composante principale (les 2/3), ce qui veut dire qu'elles sont reliées par au moins une localité d'origine. Pratiquement toutes les références dont AM déclare connaître le ou les auteur(s) se trouvent dans cette partie du réseau. Le tiers restant est composé de références dont AM ne connaît généralement pas du tout les auteurs, en particulier les références liées aux techniques et à la méthode utilisée (en bas à droite de la représentation). En dehors de la composante principale, il n'y a que les auteurs des deux références signées à Stanford dont elle déclare connaître les auteurs : l'un de nom (A. Kornberg, déjà évoqué), et les autres, parmi lesquels P. Hanawalt, parce qu'elle les a déjà rencontrés (dyade bleu en haut à droite de la représentation). Elle sait même que l'un d'entre eux est maintenant en Angleterre. Ces auteurs appartiennent à son domaine de recherche c'est-à-dire « la réparation et la transcription de l'ADN », ce qui n'est pas le cas des auteurs des références dites « techniques » ou « matériels et méthodes ». La référence au couple de chercheurs américains qu'elle ne connaît pas (la réfère-

rence n°4) est aussi une référence plus théorique que technique : c'est un couple dont elle apprécie le travail sur la transcription, mais elle pense qu'ils sont Anglais et croit savoir qu'ils ont changé de domaine. De même, la référence 32 (« très cité ») d'un groupe canadien de chercheurs qu'elle ne connaît pas est importante d'un point de vue théorique puisqu'elle montre que « Notre protéine se comporte comme une autre déjà trouvée. ».

Dans la composante principale, tout un ensemble de références ont été signées depuis Rotterdam. En fait, les 2/3 des références à des auteurs qu'elle déclare connaître (22 sur 33) sont des références qui comptent au moins une adresse à Rotterdam. Pour le dire autrement, AM déclare connaître au moins un auteur de la moitié des références de la publication (33 sur 63) et parmi ces 33 références « connues », 22 ont été signées à Rotterdam. Parmi les références aux travaux de Rotterdam, la plupart concernent les modèles murins et notamment les différents modèles de souris mutantes XP qui ont été mis au point à Rotterdam, parfois en collaboration avec Leyde et Bilthoven (cercle rouge). Une autre partie des travaux traite plus spécifiquement de la protéine XPB, parmi lesquels les travaux qu'AM a elle-même signés à Rotterdam (auto-citation qui se trouve dans le cercle jaune). Dans l'ensemble des publications signées à Rotterdam, il y a deux références à des auteurs plus proches d'elle, qui sont co-signataires de la publication : une référence à MF (« Elle était en post-doc en même temps que moi et maintenant elle est en Grèce. ») ; et une référence au compagnon d'AM lorsqu'il était également en post-doctorat à Rotterdam (« Le premier auteur de la 45 est mon mari. Il a insisté pour que je mette aussi la référence 44 sur laquelle se fonde sa publication »). Enfin, une partie des travaux signés à Rotterdam ont été co-signés avec l'équipe strasbourgeoise spécialisée dans la transcription, qui a été centrale au moment de la résolution de la question de recherche précédemment évoquée (cercle vert). L'un d'entre eux est « incontournable » (« on travaille en ce moment dans cette direction »).

Parmi tout cet ensemble de travaux en lien avec sa thématique de recherche et plus spécifiquement avec son inscription dans le laboratoire de Rotterdam et le réseau de collaboration privilégié de ce dernier qui inclut Leyde, Strasbourg, Pavie et Brighton, on comprend qu'il y a une partie des références qui sont citées préférentiellement à des références qui pourraient être équivalentes sur le plan scientifique mais issues d'un autre espace géographique (« je suis obligée de le citer », « plutôt lui qu'un autre »). En fait, les références qui sont caractéristiques de cette situation sont essentiellement des recensions : on peut en déduire qu'entre plusieurs recensions qui traitent du même sujet, il existe une prédisposition à citer plutôt les recensions issues de son propre cercle scientifique, et par conséquent de son espace géographique de collaboration. En dehors des travaux associés à son expérience de post-doctorat, AM fait référence à une publication signée à Paris par

une chercheuse dont elle a fait la connaissance pendant sa thèse, mais qui travaille chez l'Oréal. C'est le seul contact qui semble remonter à l'époque de sa thèse dans le réseau de référence.

D'une manière générale, les contacts anglais et américains connus apparaissent moins proches d'elle : elle les a seulement « rencontré en colloque », « souvent rencontré ». 4 références sont issues du *National Cancer Institute* (NCI) de Bethesda, dans la banlieue de Washington, mais même s'il y a de nombreux travaux sur la réparation de l'ADN dans ce centre de recherche, c'est plutôt pour l'aspect méthodologique qu'elle fait référence à des publications de cet organisme. Elle cite notamment le chercheur à l'origine du modèle sur lequel elle s'est fondée pour « marquer » (taguer) les protéines XPB. Ce même chercheur, elle l'apprendra plus tard, est celui qui a été désigné responsable scientifique de la publication. En dehors de cette référence d'autorité, il y a aussi plusieurs références qui discutent la pertinence du modèle où qui l'ont testé. Ces références sont assez dispersées géographiquement : Dallas, Ithaca, Osaka, Göttingen, Londres, Oxford, Rotterdam. Puisque cette méthode est assez nouvelle et délicate à utiliser, il semble important de bien argumenter pour en justifier le choix. Comme la revue n'impose pas un nombre limité de références, AM préfère souvent avoir recours à deux références qui vont dans le même sens au lieu d'une, ce qui permet d'« appuyer le résultat ».

Comme dans le cas des réseaux de références étudiés dans le cadre du programme [ANR Resocit](#)¹⁴⁴, la structure de ce réseau ne s'explique que partiellement par la trajectoire géographique d'AM. Bien sûr, une grande partie des références sont issues du laboratoire dans lequel elle vient de passer plusieurs années, mais il y a aussi tout un ensemble de références à des publications issues de lieux qu'elle n'a jamais visités et d'auteurs qu'elle n'a jamais rencontrés. Cela confirme un certain nombre de résultats connus sur les réseaux scientifiques et notamment le fait qu'une partie de la pratique scientifique repose sur la confiance qu'on accorde à ce qui a été scientifiquement validé et vérifié par d'autres, sans qu'on puisse ou qu'il soit question de le remettre en cause. Ce phénomène est connu, et il était déjà évoqué au début du XX^e siècle par Georg Simmel, considéré comme l'un des pères de l'analyse des réseaux sociaux, à l'occasion de ses réflexions sur les sociétés secrètes : « notre existence moderne – depuis l'économie, qui tend de plus en plus à être une économie de crédit, jusqu'aux activités scientifiques, où la plupart des chercheurs doit appliquer un nombre immense de résultats obtenus par d'autres, qui sont absolument invérifiables – repose, beaucoup plus largement qu'on ne le reconnaît habituellement, sur

¹⁴⁴ Les travaux réalisés dans le cadre de cette ANR ont été présentés à Toulouse les 4 et 5 Juin 2015 à l'occasion d'une journée d'étude organisée par B. Milard intitulée : « Citations scientifiques et réseaux sociaux ». Description du projet ANR : [http://www.agence-nationale-recherche.fr/suivi-bilan/editions-2013-et-anterieures/recherches-exploratoires-et-emergentes/blanc-generalite-et-contacts/blanc-presentation-synthetique-du-projet/?tx_lwmsuivibilan_pi2\[CODE\]=ANR-11-BSH1-0013](http://www.agence-nationale-recherche.fr/suivi-bilan/editions-2013-et-anterieures/recherches-exploratoires-et-emergentes/blanc-generalite-et-contacts/blanc-presentation-synthetique-du-projet/?tx_lwmsuivibilan_pi2[CODE]=ANR-11-BSH1-0013)

la foi en l'honnêteté des autres. »¹⁴⁵ (Simmel, 1908a, p. 16). Mais, au-delà d'offrir une démonstration de ce phénomène, notre analyse permet d'en saisir la traduction géographique : on constate effectivement que les travaux convoqués par AM sans en connaître les auteurs (pas même de nom), sont souvent des travaux qui ne relèvent pas de l'espace cognitif de sa question de recherche (ils sont davantage en relation avec les techniques et la méthode utilisée). À l'inverse, les travaux qui sont les plus proches de sa question de recherche sont aussi les plus proches de l'espace géographique de collaboration dans lequel elle évolue. Dans cet espace-ci, la vérification est beaucoup plus la norme, et AM sait que les équipes qui travaillent précisément sur la même question qu'elle vont chercher à reproduire son expérience et à l'étendre. On peut enfin remarquer que les références d'autorité aux pionniers de sa spécialité de recherche restent extrêmement limitées – il y a les deux références aux auteurs de Stanford et une référence à une review de D. Bootsma qui date de 1993 (« pas trop récent ») – ce qui suffit à positionner l'article, qui par ailleurs est un article assez tourné vers la technique parce qu'il s'agit de rendre compte d'une avancée technique dans une revue qui met justement l'accent sur le caractère innovant des techniques utilisées et des résultats trouvés (*Plos ONE*).

Au-delà de pouvoir être liées entre elles par une localisation géographique et analysées suivant cette dimension, les références de la publication permettent de replacer la publication dans un espace cognitif, et à travers elle AM et ses co-auteurs. En effet, il est possible, en partant de la publication, de délimiter un corpus de publications qui citent les mêmes références que la publication. Ce corpus peut être considéré comme « l'univers de référence » de la publication (Milard, *op. cit.*). Ici nous avons reconstitué cet univers à partir des données du *Web of Science* pour en étudier l'inscription géographique, et en particulier, l'espace de collaboration dont il est le produit. À cette fin, nous avons considéré seulement l'ensemble des publications qui ont cité au moins 3 références en commun avec l'article d'AM. Il y en a 360, qui ont été publiées entre 1994 et 2012. Nous remarquons ainsi que l'inscription géographique de ce dernier corpus recoupe dans une certaine mesure l'espace de la question de recherche sur le lien entre réparation et transcription de l'ADN.

¹⁴⁵ Le fait qu'une partie importante de la science repose sur la confiance qu'on accorde aux travaux des autres indique que la pratique de la science n'est pas aussi conditionnée par l'acquisition du savoir tacite des autres chercheurs qu'on le raconte souvent. Quand bien même on chercherait à ce que tous les scientifiques partagent leurs savoirs tacites, ils n'y arriveraient pas car cette dimension du savoir n'est ni entièrement consciente, ni entièrement inscrite dans la mémoire de son détenteur. Ainsi, lors d'une ethnographie de laboratoire, « en amenant les chercheurs à la fouille de leur propre documentation, Sébastien Plutniak et Dimitri Aguera ont constaté, qu'ils [les chercheurs] peuvent, eux-mêmes, ne plus s'y retrouver. L'auteur se dissout ici à lui-même, en la perte du savoir tacite qu'il avait incorporé. Les cahiers sont devenus illisibles à leur propre scripteur. » (Plutniak & Aguera, 2013).

Nous avons agi pour ce corpus de la même manière que pour le corpus extrait à partir des mots-clés de la question de recherche, c'est-à-dire que nous en avons extrait un réseau de co-signature entre lieux. La seule différence est de pas avoir procédé à l'agrégation des localités au niveau des agglomérations scientifiques, ce qui n'était pas indispensable compte tenu du faible nombre de publications considérées ici (360). Le résultat montre l'espace géographique des collaborations qui sous-tendent l'univers de référence de la publication d'AM. Seule la composante principale du réseau est représentée (Figure 91).

Réseau des co-signatures de l'univers de référence de la publication (au moins 3 références communes) en jan. 2013

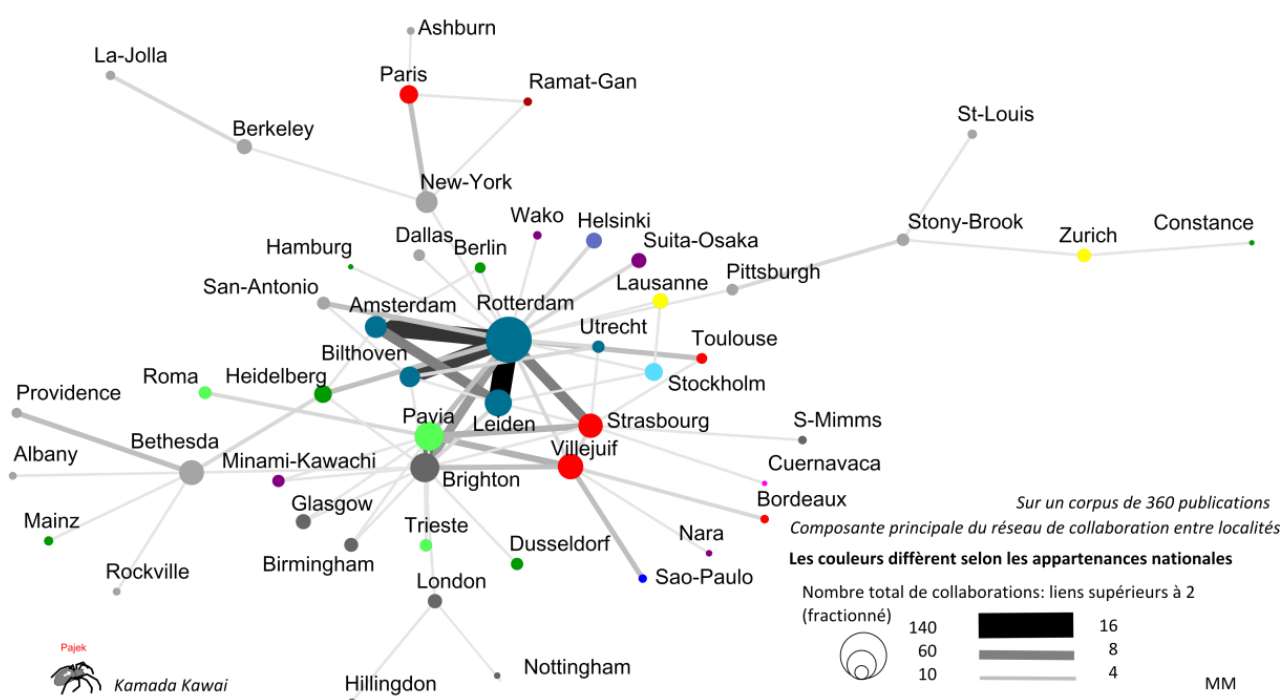


Figure 91 – Réseau de co-publication des 360 publications citant au moins 3 références communes à l'article scientifique commenté en entretien. Source : Web of Science

Ce résultat est parfaitement cohérent, compte tenu du contexte de production de la publication (à Rotterdam) mais aussi de la connaissance que nous avons de la géographie de la question de recherche dans laquelle s'inscrit cette publication. Majoritairement, les publications qui ont cité les mêmes références que la publication d'AM ont été produites aux Pays-Bas par l'ensemble des équipes spécialisées dans la question de recherche sur le lien entre transcription et réparation de l'ADN. Ces publications, sont aussi le produit de collaborations régulières avec le cercle privilégié de la « Dutch army » : Pavia, Brighton, Strasbourg. On repère aussi quelques éléments moins attendus comme le nombre de pu-

blications issues de Villejuif alors qu'AM a peu cité cette équipe de l'Institut Gustave Roussy, qui entretient des relations privilégiées avec l'Amérique Latine sur cette question de recherche : en particulier avec São Paulo, où un ancien post-doctorant à Villejuif est retourné vivre pour monter son équipe.

Au-delà, comme l'univers de référence comprend des publications qui sont parues après 2007 et même après la création de l'équipe d'AM, on remarque l'intégration de Toulouse au réseau et son lien de co-publication avec Rotterdam, alors que la ville n'apparaissait pas encore sur la question de recherche avant l'arrivée d'AM. Dans la même logique, il aurait été normal de trouver la nouvelle équipe de MF à Athènes sur la représentation, mais il n'en est rien ; l'hypothèse la plus probable pour justifier cette absence étant qu'elle se soit lancée, depuis la Grèce, sur une question de recherche un peu différente. Enfin, on peut noter qu'indépendamment de la cohésion néerlandaise sur le sujet, il n'y a pas d'autres cohésions nationales remarquables. Les équipes appartenant au même pays et qui ont en commun de citer les mêmes références qu'AM n'ont pas tendance à collaborer entre elles. Cela est sans doute lié à la part prépondérante des travaux néerlandais qui se trouvent parmi les références d'AM. Les publications qui appartiennent à ce corpus y figurent en grande partie parce qu'elles ont cité des travaux hollandais. Pour le dire autrement, le dénominateur commun de la plupart des publications présentes dans cet univers de référence est d'avoir cité, ou d'avoir collaboré avec une équipe des Pays-Bas.

La représentation montre qu'en dépit de son caractère très européen, cette aire d'influence s'étend en Israël, aux États-Unis, au Japon, au Mexique et au Brésil, mais pas en Chine ; ce qui confirme le fait que cette question de recherche reste investie essentiellement par des pays engagés depuis longtemps dans le domaine de la réparation de l'ADN. Comme, il existe une division des tâches, y compris parmi les équipes qui travaillent sur le lien entre réparation et transcription de l'ADN, et que par exemple le travail expérimental sur les souris mutantes XP est une spécialité de Rotterdam, il est probable qu'on n'observe ici uniquement les équipes qui travaillent sur ce même axe de la question de recherche ; d'où le fait que l'aire géographique des collaborations ne recoupe que partiellement celle de la question de recherche à l'échelle mondiale. Il ne s'agit bien sûr que d'hypothèses qui mériteraient de plus amples vérifications pour être validées.

On pourrait aller plus loin dans cette analyse et, par exemple, opérer une distinction entre les publications parues avant et après la publication d'AM pour mieux saisir les évolutions en cours et la place de la publication d'AM dans ce contexte présent. Aussi, il aurait été possible de regarder plus précisément la structure de l'univers de références à travers les liens de co-citations entre les publications de cet univers : les publications les plus connectées étant alors celles qui partagent le plus de références, et les isolats celles

dont la ou les références ne sont partagées avec aucune autre des publications de l'univers (ces pistes sont exploitées par Milard, 2011b, *op. cit.*). Mais, avec l'idée d'aller un peu au-delà dans la compréhension du positionnement géographique de la publication d'AM, nous avons choisi d'explorer l'espace de collaboration de l'univers de référence au regard de l'espace de collaboration de la question de recherche étudiée dans la première section du chapitre et dans laquelle s'inscrit la publication d'AM.

Dans l'ensemble, cette analyse à plusieurs niveaux scientifiques de la géographie d'un domaine de recherche est l'occasion d'envisager plus finement l'inscription spatiale des activités scientifiques. Comme Mullins l'a montré, les phénomènes de développement de nouvelles questions et d'intégration de nouveaux membres se produisent à différents moments de l'histoire de la spécialité, et pas seulement à son stade paradigmatique, c'est-à-dire qu'à mesure qu'elle évolue, on retrouve de façon fractale les mêmes processus d'évolution mais à l'échelle de sous-ensembles de recherches. Aussi, même au niveau d'une carrière on peut repérer différents stades dans la structure du réseau sociocognitif qui peuvent traduire des changements de trajectoires, mais aussi être le fruit de rencontres, de lectures qui influencent le positionnement relatif des chercheurs dans l'espace sociocognitif et par là même géographique. Au moment où AM a publié l'article qu'elle a choisi de commenter, son positionnement révèle son appartenance à un cercle de collaboration, celui de l'équipe qu'elle est en train de former et celui de l'équipe néerlandaise dont elle est issue. Les références de sa publication traduisent bien l'influence qu'exercent sur son travail les travaux du cercle plus élargi des collaborateurs réguliers de son équipe d'origine. Ce cercle est un collectif cohésif à la fois au niveau national (« *Dutch army* ») et européen (un réseau de collaboration soutenu par des institutions européennes). Enfin, la géographie des références de sa publication et de l'univers de référence de sa publication montre comment son positionnement cognitif, contrairement à ses pratiques tangibles de recherche, dépasse parfois, en termes d'aire géographique, son positionnement social, puisqu'elle a connaissance et se réfère à des travaux qui appartiennent à des aires de collaborations extérieures à l'Union Européenne, et qui traduisent, notamment, la qualité mondiale de sa communauté de recherche et son inscription dans un héritage qui remonte à l'après Seconde Guerre mondiale.

CONCLUSION

Il existe de nombreuses sources sur lesquelles s'appuyer pour localiser les activités de production scientifique et repérer les liens scientifiques de collaboration, de mobilité, ou d'influence qui s'établissent entre les lieux d'exercice de l'activité scientifique à l'échelle mondiale. Après plusieurs décennies au cours desquelles le *Science Citation Index* (SCI) occupait une position inégalée comme outil de recherche d'information bibliographique et comme outil d'analyse de la répartition et des dynamiques mondiales de production scientifique, d'autres index bibliographiques se sont hissés au rang de sérieux concurrents durant la dernière décennie. La visibilité et l'impact de la littérature recensée, le caractère exhaustif et la qualité de la couverture géographique sont les éléments qui animent la concurrence que se font depuis l'ISI Thomson Reuters (pour le *Web of Science*), Elsevier (pour *Scopus*) ou encore Google (pour *Google Scholar*). À côté de ces géants, nombres de sources bibliographiques alternatives co-existent qui sont centrées sur des disciplines en particulier, des régions du monde bien délimitées, ou encore des contenus spécifiques.

Pour déterminer sur quelle source s'appuyer, ou éventuellement sur quelle combinaison de sources, il paraît important d'avoir une bonne connaissance des biais inhérents à chacune d'entre elles, ainsi que de leurs évolutions. S'il s'agit d'un axe de recherche qui importe naturellement aux bibliomètres, le contrôle des biais géographiques de ces sources est un point qu'il ne faudrait pas oublier de mettre à l'agenda de la géographie des activités scientifiques. Cette géographie, au nom de laquelle nous tenons (et d'autres avec nous) le rôle d'éclaireur, trouve, entre autres raisons d'être, le besoin commun aux chercheurs et aux administrateurs de la science d'avoir une image juste de l'étendue de la production scientifique à l'échelle du monde. Non content de pouvoir ainsi tenir le rôle de cartographe des sciences, le géographe est aussi celui qui serait en mesure, au côté des autres spécialistes des sciences, de dévoiler la mécanique sous-jacente allant des lieux de l'écriture aux lieux de l'évaluation, puis des lieux d'édition aux lieux de l'indexation, et enfin aux lieux de diffusion du savoir scientifique. Débordant la géographie des publications scientifiques ainsi mise à jour, il resterait encore à étudier tout un pan du travail de recherche qui ne se réduit pas à la publication d'écrits, tel que l'expérimentation ou le terrain et dont l'inscription géographique ne saurait, en l'espèce, être entièrement résumée par l'adresse professionnelle des auteurs.

Le programme de la géographie des activités scientifiques est vaste et ne s'assimile pas à celui de la géographie des connaissances, ni non plus à ceux de la géographie de l'innovation et de l'enseignement supérieur. Dans cette perspective, la localisation au niveau urbain de l'ensemble des publications scientifiques indexées dans le *Web of Science* (WoS) au cours de la décennie précédente est l'objectif central qui a été atteint par notre équipe de recherche, auquel cette thèse a contribué et sur lequel elle s'est appuyée.

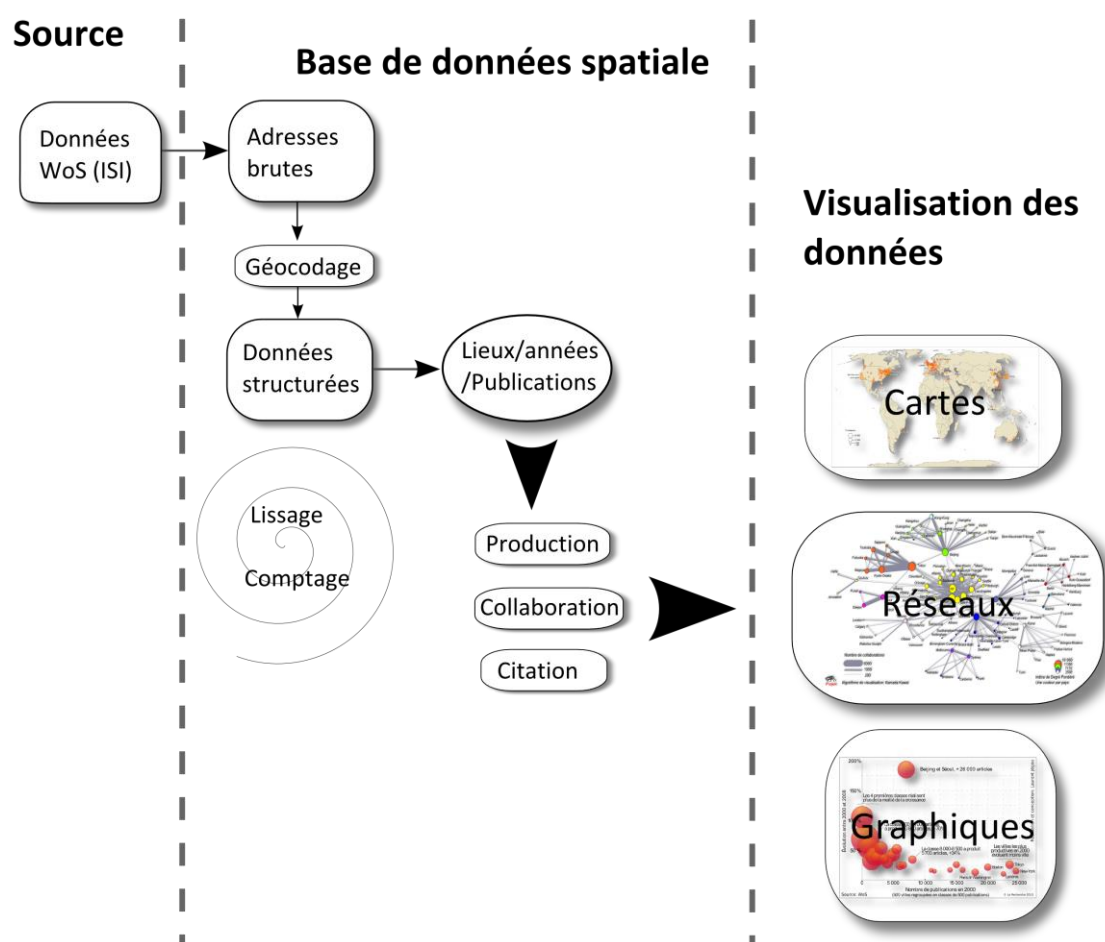
Parallèlement au programme de recherche « Géoscience », la thèse a plus précisément investi la question des socles théoriques existants pour que se développe une géographie des sciences ; des biais et de l'histoire de la source bibliographique mobilisée pour localiser l'activité scientifique ; de l'influence des choix méthodologiques sur les observations (méthodes de comptage, périmètres d'agrégation) ; de l'évolution de la structure des réseaux de collaborations toutes disciplines confondues, mais aussi au niveau d'une question de recherche en biologie moléculaire ; de l'émergence d'un réseau de collaboration entre lieux à l'échelle mondiale dans le domaine de la réparation de l'ADN, ainsi que de l'inscription de trajectoires individuelles dans ce réseau de lieux, considérant à la fois l'influence de ces trajectoires dans la structuration d'un espace communautaire et le caractère structurant des institutions participant à l'organiser.

Ainsi, le détour par l'histoire de la source mobilisée a permis de rappeler que le WoS est une source de référence pour les bibliomètres depuis les années 1960 et donne à voir un aspect de la production scientifique mondiale : celui des revues se conformant aux « standards internationaux » dictés par la science américaine. Dans cet ensemble, l'index des sciences formelles (le SCI Expanded) et dans cet index un certain type de documents (les articles, recensions et lettres) ont été retenus par notre équipe pour une analyse globale de la répartition de la production scientifique et de son organisation à plusieurs niveaux (pays, grandes aires de collaboration) à l'échelle mondiale. Ce choix a été fait pour être plus représentatif de la réalité de l'activité scientifique, sachant que le WoS dispose d'une moins bonne représentativité dans le domaine des sciences humaines et sociales et d'une moindre influence sur les pratiques de publication qui en relèvent.

Cependant, l'intégralité de l'information géographique intégrée dans la base de données a été considérée pour mettre au point la méthode de traitement des données consistant à : coder les adresses de publication (leur affecter des coordonnées géographiques), les agréger au niveau d'entités urbaines comparables à l'échelle mondiale (les agglomérations scientifiques fondées sur le respect d'un critère de densité de population pour les plus grandes d'entre elles et de distance géographique pour les moins publiantes), en extraire des stocks de publications par lieu et des matrices de relations entre lieux (à partir d'un comptage entier fractionné des contributions et des liens de co-signatures au niveau des agglomérations), et les représenter sous la forme de cartes, de réseaux et de graphiques ou tableaux statistiques. Cette chaîne procédurale ou démarche de travail définie en équipe se résume sous la forme du schéma récapitulatif suivant (Figure 92). Ce protocole, défini à partir de l'ensemble des données extraites du WoS, a été également prévu pour s'adapter à des sous-ensembles du WoS ainsi qu'à des ensembles de données bibliographiques issus d'autres sources bibliographiques.

La reconstitution de la carte de l'émergence du domaine de la réparation de l'ADN entre 1965 et 1975 en témoigne, qui a été réalisée à partir de données extraites de la base *Scopus*. Ainsi, plusieurs ensembles de données bibliographiques ont été mobilisés pour se rapprocher des contenus d'un domaine de recherche en biomédecine et, au-delà, des pratiques scientifiques et des trajectoires individuelles des chercheurs. Le recours aux données du WoS et de *Scopus* dans ce cadre a fait ressortir d'autres limites liées aux sources comme la piètre qualité d'indexation des notices de publication antérieures à 1990 ou l'impossibilité de se fonder seulement sur les marqueurs de publication pour suivre la mobilité des scientifiques.

Le traitement géographique de données bibliographiques



Conception et réalisation: Laurent Jégou et Marion Maisonobe

Protocole de travail établi dans le cadre des projets géoscience->netscience

Figure 92 – Schéma récapitulant la chaîne de traitement spatial des données bibliographiques

Malgré les limites intrinsèques à la mobilisation de données bibliographiques, la thèse montre comment ces dernières peuvent être abordées en premier lieu, comme témoins de la localisation des activités de recherche à l'époque contemporaine et dans un second temps, comme traces historiques des liens entre lieux et individus pour des études de géographie socio-historique à l'échelle mondiale.

Interrogeant l'évolution récente de la répartition des publications et de la structure du réseau mondial des collaborations interurbaines, les analyses de scientométrie spatiale menées au niveau des agglomérations urbaines sur l'ensemble des articles, recensions et lettres indexés dans le SCI Expanded ont fait ressortir trois faits saillants.

Premièrement, la production scientifique s'est dispersée géographiquement en se développant, c'est-à-dire que l'augmentation globale de la production entre 2000 et 2007 a d'abord profité aux espaces secondaires ou périphériques de production. Ces derniers ont bénéficié de taux de croissance relatifs beaucoup plus élevés que les espaces d'activité historiquement centraux, tels que les capitales scientifiques à l'échelle des États, ou l'Occident à l'échelle mondiale. Autrement dit, nous assistons à un rééquilibrage à plusieurs échelons de la répartition des activités de production scientifique, traditionnellement extrêmement concentrées.

Deuxièmement, le réseau mondial des collaborations scientifiques interurbaines s'est densifié globalement, mais avec des différences entre les pays, selon qu'ils traversaient une phase de forte déconcentration de leur production (Turquie, Chine, Iran...) ou bien une stagnation, voire une légère concentration entre 2000 et 2007 (États-Unis, Afrique du Sud, Canada, Royaume-Uni, Japon...). En particulier, la croissance des liens intra-nationaux, ou plus précisément la préférence pour les collaborations domestiques, a été relativement plus forte pour les agglomérations appartenant au premier type de pays. En dehors du fait que la géographie des collaborations suit la géographie de la production, cette observation est apparue caractéristique du rôle de la cohésion des systèmes nationaux de recherche dans la structuration des réseaux de collaborations interurbaines, puisque dans les pays où la production s'est le plus diffusée, les collaborations domestiques se sont renforcées au détriment des liens avec l'international.

Troisièmement, tandis que la part des publications à une seule adresse a diminué en tous lieux, l'augmentation des collaborations est intervenue à différents niveaux : au niveau intra-urbain, au niveau national et au niveau des grandes aires de collaborations mondiales. L'augmentation des collaborations a aussi bénéficié aux connexions entre les grandes aires de collaboration, et notamment aux relations entre espaces périphériques. En fait, selon l'échelle géographique à laquelle on s'est placé et la discipline, la croissance des collaborations est apparue plus ou moins marquée, de sorte que, contrairement à l'image d'un espace mondial traversé par des tendances homogènes et partagées par tous

dans les mêmes proportions, nous avons affaire à des phénomènes de rééquilibrage, au renforcement de la cohésion de certains espaces au détriment d'autres espaces, au maintien du caractère structurant des systèmes régionaux, nationaux, et de certaines grandes aires de collaboration ; le tout accompagné d'une moindre hégémonie des États-Unis et plus globalement de l'Occident dans la production scientifique mondiale et le système de collaboration.

Globalement, les pays ayant connu une période de décentralisation dans les dernières décennies ont été concernés par une tendance plus grande au polycentrisme qui s'explique par la répartition plus équitable des activités de production scientifique en leur sein. Dès lors, contrairement à l'idée d'un inévitable renforcement des écarts entre agglomérations lié à la mondialisation, les plus grandes étant vouées à participer aux échanges mondiaux quand les autres devraient se cantonner à évoluer dans l'ombre des premières, nous avons repéré que l'appartenance nationale était, dans bien des cas, plus déterminante que la taille des agglomérations pour expliquer l'évolution de leur comportement de production. Il y a ainsi des tendances nationales à la stagnation pour les agglomérations de Russie, à la croissance pour celles de Chine et, enfin, au polycentrisme (en matière de collaborations) pour celles qui se trouvent dans des pays en voie de déconcentration. Bien que rarement aussi soutenue que la progression des liens intra-nationaux, celle des liens internationaux s'est déroulée bien souvent en faveur de la cohésion de grandes aires régionales de collaboration, et seul l'espace états-unien a été concerné par une baisse de sa cohésion interne en faveur de ses échanges avec l'extérieur. La mondialisation ou le renforcement des liens entre les différents espaces géographiques de la planète n'a pas été invalidé par nos résultats. Simplement, contrairement à ce qu'affirment des travaux menés uniquement sur un sous-ensemble de grandes villes ou au niveau des pays, cette mondialisation n'a pas lieu au détriment de la poursuite d'autres processus comme la nationalisation, la régionalisation et la continentalisation des échanges entre agglomérations urbaines.

Ces grandes tendances sont le résultat de la somme des dynamiques de production scientifique qui interviennent au niveau des collectifs de recherche. Agrégées toutes ensembles ces tendances résultent du fait que les questions de recherche et les terminologies tendent à se diffuser dans l'espace géographique, et de nouvelles équipes peuvent, en se connectant aux premières, venir y contribuer. Ces processus de diffusion sont facilités par les échanges de savoir-faire par le biais de la mobilité des scientifiques, mais aussi par la circulation de ressources plus matérielles. L'inégale répartition des ressources nécessaires aux travaux de recherche contraint la géographie des activités scientifiques, ce qui explique la plus ou moins grande concentration des activités dans certains lieux.

Si elles peuvent se produire à différents niveaux, les interactions, et en particulier les collaborations, sont plutôt locales ou nationales car les dispositifs de coordination (la création de postes, l'organisation de jury de thèse, la distribution des ressources financières et matérielles...) mais aussi la société en général (les systèmes éducatifs, les réseaux de transport et de télécommunication, les réseaux de distribution de produits culturels...) sont d'abord organisés à ces niveaux-là. Sans qu'il s'agisse d'une fatalité, la probabilité d'avoir des ressources à partager et échanger avec nos plus « proches voisins » est élevée, d'autant plus s'ils sont nos « collègues » sur le plan institutionnel. Mais, cela n'empêche pas les interactions scientifiques de pouvoir se développer au niveau international, parce qu'il existe, dans le milieu de la recherche, depuis bien longtemps, une volonté d'organisation à l'échelle mondiale. Ces formes d'organisation existent au niveau de chaque discipline par le biais d'associations ou de sociétés savantes, et parfois très explicitement au niveau de certaines « communautés » de recherche comme la communauté de la réparation de l'ADN.

Ce collectif dont les origines remontent aux années 1960 est né de l'intérêt partagé par un ensemble de biophysiciens, radiobiologistes, photobiologistes, biologistes cellulaires et moléculaires, ou encore généticiens pour les réponses intracellulaires aux lésions d'ADN. Son organisation a été facilitée par les moyens dont disposaient les chercheurs issus des recherches sur les radiations au sortir de la guerre et, notamment, les équipes américaines pour accueillir et former des chercheurs, organiser des rencontres, soutenir le développement des groupes européens et japonais. Dans ce domaine, des sous-questions se sont rapidement distinguées et une répartition des tâches s'est mise en place entre les groupes de recherche selon le mécanisme de réparation de l'ADN et selon le type d'organisme vivant étudié. Les équipes néerlandaises et, en particulier, celle dirigée par D. Bootsma à Rotterdam ont ainsi fait preuve d'une grande expertise dans la compréhension du mécanisme de réparation de l'ADN par excision de nucléotide (NER). Cette maîtrise s'appuyait sur une compétence dans l'analyse et la différenciation des cellules de patients atteints de *Xeroderma Pigmentosum* (XP), qui petit à petit a amené le groupe de recherche à faire d'importantes découvertes. Au début des années 1990, les équipes de Rotterdam et Leyde, avec l'aide d'une équipe strasbourgeoise spécialisée dans la transcription ont ainsi trouvé la preuve de l'existence d'un lien entre réparation et transcription de l'ADN à travers le facteur TFIIH, ce qui a donné lieu à de nombreux travaux. Bien que cette avancée ait été connue et évoquée par un nombre grandissant d'équipes à l'échelle mondiale dans les années 2000, l'équipe de Rotterdam a joué un grand rôle dans l'organisation du collectif travaillant précisément sur cette question à l'échelle européenne. La carrière européenne de nombre de post-doctorants de Rotterdam en atteste et, notamment, la trajectoire d'une chercheuse ayant obtenu la possibilité de créer une

nouvelle équipe sur cette question à Toulouse à la suite de son post-doctorat, équipe actuellement en train de déménager pour Lyon.

Avant ce phénomène de diffusion dont les logiques révèlent l'existence de niveaux d'organisation locaux, nationaux et continentaux, la découverte faite à Rotterdam a d'abord suscité, au début des années 1990, l'activation plus exceptionnelle de liens entre grands espaces mondiaux. Ce phénomène nous a permis de supposer que les connexions les plus rares entre grands espaces de collaboration étaient plus volontiers activées à certains moments stratégiques pour le progrès scientifique. Dans le même temps, au niveau des trajectoires de recherches, les liens extérieurs à l'espace géographique privilégié de collaboration des chercheurs se manifestent par les références et le processus de diffusion de leurs recherches dans des revues à « impact international », mais l'espace concret de l'activité demeure peu étendu et réparti entre un nombre plus limité de lieux. En ce sens, notre travail montre qu'il est possible de saisir la traduction spatiale des cycles ou étapes caractéristiques de l'organisation sociale des pratiques de recherche au sein de collectifs plus ou moins institutionnalisés (de la question de recherche à la discipline en passant par la spécialité). La connaissance que possèdent les sociologues et historiens des sciences de ces dynamiques peut ainsi être complétée par une maîtrise et une capacité à en mesurer les manifestations dans l'espace géographique et, parallèlement, les géographes intéressés par la compréhension des dynamiques humaines à plusieurs échelles peuvent se nourrir de cette meilleure connaissance des logiques d'organisation et d'institutionnalisation du monde scientifique contemporain pour mieux interpréter les cartes de l'activité scientifique, et avec elles, les cartes des activités humaines en général.

Si l'aspiration à la « visibilisation » du travail scientifique a été partagée par un nombre croissant de lieux d'activité au cours de ces dernières décennies, suggérant, même s'il faudrait le nuancer, l'existence d'une motivation commune à un nombre croissant de scientifiques d'intégrer le grand « collège » de la science mondiale, les frontières nationales ont conservé une dimension structurante du fait, notamment, des particularismes des systèmes de recherche. Il y a bien sûr des imitations, des échanges mais les systèmes nationaux restent bien différenciés. Pour des raisons historiques, il existe même une différenciation entre les différents espaces urbains d'un même pays, qu'il serait vain de chercher à éliminer. En effet, comme le reconnaissait Olivier Dollfus dans son ouvrage sur la mondialisation : « Par suite de toutes les rugosités à la fois spatiales et sociales, la concurrence parfaite entre les lieux n'existe pas, les différences se maintiennent à travers les potentialités des lieux. » (Dollfus, 1997, p. 111). Ces différences sont constitutives de la richesse et de l'hétérogénéité de la science mondiale et sont, bien sûr, au cœur du projet de géographie des sciences dont cette thèse se fait l'écho. Pour rebondir sur la référence littéraire qui ouvrait le premier chapitre (Lodge, 1975), même en donnant des noms ima-

ginaires aux villes universitaires entre lesquelles le romancier David Lodge fait voyager les deux protagonistes principaux de son roman *Changement de Décor*, on reconnaît sans trop de difficulté que la ville américaine très ensoleillée abritant une université réputée et agitée par d'importants soulèvements étudiants en 1968 n'est autre que Berkeley, tandis que la ville anglaise de tradition industrielle à brique rouge et au climat peu clément se trouve quelque part dans les Midlands (les connaisseurs déduiront des détails et de la carrière de David Lodge qu'il s'agit d'un avatar de Birmingham).

Si les mécanismes et le fonctionnement de la recherche nous viennent du siècle dernier et même parfois remontent à de plus anciens héritages, il est indéniable que les quantités d'établissements, de scientifiques, de problèmes de recherche, de publications qui y sont associées se sont démultipliées ce qui, avec la multiplication des modes et des échelons de coordination, donne le sentiment d'une complexification et d'une plus faible lisibilité du paysage global de l'activité scientifique. Ce sont ces phénomènes concrets qui rendent louable et impérative la nécessité de mieux décrire les particularités et les variations qui traversent les lieux de science et participent de leur positionnement au sein de réseaux de lieux. Comme toujours pour les secteurs qui aspirent à un certain mondialisme, ou même en l'occurrence à l'universalisme, la standardisation des pratiques et des idées est redoutée. Cette crainte n'est pas infondée étant donné les rapports de domination qui, par exemple, valorisent l'usage de l'anglais sur celui de toutes les autres langues pour le travail scientifique. À ce titre, les résistances et les alternatives au système américanisé, sur lequel nos analyses se sont fondées, mériteraient bien sûr d'être considérées, ce qui justifierait de se rapprocher encore davantage de la littérature portant sur la recherche dans les espaces périphériques. Ces affaires de domination, de centres et de périphéries, de migrations parfois, d'accès aux techniques et réseaux de communication, de pouvoir aussi, font de ce champ de recherche en géographie un domaine capable de dialoguer avec tout un ensemble d'autres axes bien plus traditionnels pour la discipline. Pour ouvrir les perspectives, nous allons tâcher d'esquisser ici quelques pistes d'interfaces qui nous semblent prometteuses. Aussi, on verra qu'une géographie des sciences fondée sur des emprunts aux études sociales des sciences a tout intérêt à progresser à l'unisson avec ces dernières, qui n'ont pas fini de bénéficier de l'apport et des approches venues de nouvelles disciplines.

Bien qu'inévitablement liée à la géographie de l'innovation et de l'enseignement supérieur et compte tenu qu'elle en est, en quelque sorte, l'émanation, notre thèse est que la géographie des activités scientifiques n'en possède pas moins ses propres logiques qui se distinguent, au moins partiellement, des logiques économiques associées aux dynamiques d'innovation et des logiques de formation du public étudiant associées aux dynamiques d'enseignement. Néanmoins une partie de ce public nourrit la recherche tandis

que cette dernière trouve aussi sa raison d'être et un certain nombre de ses problèmes auprès de l'industrie et plus généralement des mondes de l'entreprise.

Comme les géographies de l'enseignement et de l'innovation, la géographie des sciences est déterminée par des mobilités et des flux de toutes sortes. Parmi ceux-ci, on compte les flux migratoires qui intéressent de près la géographie des « Suds » et aux côtés de laquelle, toute une géographie des mobilités savantes pourraient être programmée. Les données manquent sur cette mobilité non pas juste du personnel « qualifié » mais aussi « scientifique » et en particulier sur la compréhension du lien entre cette dernière et la géographie des objets et problèmes de recherche. Comme en témoigne un récent travail de l'OCDE (Appelt, van Beuzekom, Galindo-Rueda, & de Pinho, 2015), les données bibliométriques apparaissent comme un moyen possible de contourner ce manque. Mais les limites de cette approche (Chapitre 10), tout comme ses apports gagneraient à être considérés de près par des géographes familiers des flux migratoires en général, et donc capables de qualifier l'information disponible sur les flux de chercheurs et la mettre en relation avec les dynamiques d'ensemble du système mondial.

Davantage encore que la géographie de l'enseignement et de l'innovation, la géographie des sciences est déterminée par des enjeux géopolitiques, puisqu'on prête à la science un rôle décisif en temps de guerre comme en temps de paix dans l'organisation des relations internationales et le contrôle des territoires. En servant un enjeu apparemment détaché des intérêts politiques, la science est un instrument décisif dans l'administration de ce qu'il est désormais commun d'appeler le « *soft power* ». La visibilité de la Chine à la fois dans le domaine des sciences et du sport est essentielle dans l'affirmation de son pouvoir sur la « scène internationale », qui mériterait de plus amples approfondissements. À ce titre, la géographie des sciences est en mesure de révéler des alliances, des changements dans l'ordre établi, mais aussi de mieux voir où se portent les intérêts nationaux à travers les terrains et objets de recherche qui bénéficient d'investissements publics. Aussi, les volontés d'organisation à des niveaux supranationaux comme le niveau européen sont le fruit de stratégies territoriales dont l'enjeu n'est pas seulement scientifique, ni non plus seulement économique comme le suggère la géographie de l'innovation, mais aussi politique. La tradition en géographie politique d'analyser les relations entre États et organisations internationales (Beauguitte, 2011) offre un axe intéressant pour interpréter et analyser les grandes tendances mondiales qui traversent la géographie des activités scientifiques.

Comme pour un très grand nombre de domaines dont l'enseignement, l'innovation, mais aussi la culture et la santé, l'organisation des activités scientifiques est traversée par les mutations associées à l'adoption et l'exploitation des nouvelles techniques d'information et de communication. Or ces mutations ont des effets sur la géographie traditionnelle des domaines qu'elles touchent (Marchandise, 2013). En particulier, la

numérisation des revues et l'accès facilité à l'information scientifique *via* Internet sont susceptibles d'atténuer les inégalités d'accès aux ressources documentaires et aux fenêtres de visibilité. Cependant, les spécialistes du numérique et des nouvelles technologies ont montré dans quelle mesure les nouveautés pouvaient reconfigurer des secteurs mais aussi favoriser l'apparition de nouvelles inégalités. En particulier, tout un ensemble de travaux de géographie portent sur les manifestations de ces mutations dans l'espace géographique et leurs effets sur la cohésion et les liens entre territoires. Ici encore, les enjeux spécifiques à la géographie de l'activité scientifique mériteraient d'être appréhendés et mis en relation avec le savoir dont on dispose en géographie des télécommunications et du numérique.

De même, la géographie de la santé est étroitement liée à la géographie des sciences, au moins biomédicales, puisque la répartition des maladies est en mesure de déterminer ou contraindre la répartition des recherches, en témoigne le cas de l'accès aux cellules de patients XP pour étudier le NER dans le domaine de la réparation de l'ADN, mais aussi parce que la répartition de la population et les particularités des différents systèmes nationaux de santé ont une influence sur la répartition et le fonctionnement des hôpitaux. De ce fait, ces géographies gagneraient sans doute, elles aussi, à dialoguer ensemble. Ce phénomène de lien entre l'application des recherches et leur répartition est aussi susceptible de se retrouver dans d'autres secteurs dont les secteurs industriels, qui ont pour le moment été au centre des études à caractère géographique réfléchissant aux « transferts ». Mais les transferts ont jusqu'ici été analysés dans le sens allant de la recherche à l'industrie plutôt que dans le sens inverse, c'est-à-dire en envisageant que la présence de tel ou tel secteur industriel puisse expliquer la présence et le développement de tel ou tel axe de recherche en un lieu donné.

Tandis que la sociologie des institutions scientifiques, telle qu'elle avait été esquissée par Joseph Ben-David dans les années 1960-1970, a refait surface en lien avec la géographie de l'innovation et de l'enseignement supérieur, l'émergence de la géographie des sciences est par ailleurs associée à tout un courant de littérature, plutôt anglo-saxon, qui traite de la géographie des connaissances, des arts et du savoir en général. Cette approche porte avec elle une dimension historique et épistémologique qui l'apparente à la géographie historique et parfois même à l'histoire de la géographie (Cuyala, 2014). Ce domaine dont l'apparition a été justifiée par la rhétorique du « *spatial turn* » est intimement connecté à celui des *science studies*. À ce titre, de nombreuses passerelles peuvent se mettre en place dans la mesure où l'approche géographique est susceptible d'éclairer un grand nombre de questions qui animent le champ des études sur la connaissance et le savoir.

La traduction spatiale du phénomène d'émergence d'un collectif de recherche, de résolution d'une question de recherche, d'institutionnalisation d'une discipline ne sont que des exemples parmi d'autres des nombreuses entrées possibles dans cette perspective.

Tout ce qui relève des effets liés à l'instrumentation, à la circulation des spécimens pourrait, par exemple, être envisagé sous un angle géographique. Nombre de travaux en histoire sociale des sciences tendent actuellement à intégrer une dimension géographique, et les géographes ont tout intérêt à les y accompagner, pour pouvoir à la fois nourrir la géographie des sciences de cas particuliers et lui apporter une dimension historique. Avec la sociologie des sciences aussi, il existe évidemment des échanges que cette thèse a partiellement permis d'envisager à travers l'importation des analyses de trajectoires de chercheurs. Mais il serait aussi possible d'établir des liens avec les travaux de sociologie des sciences soucieux des modes de régulation de l'activité scientifique et, notamment, des effets sur le développement scientifique de l'application de différentes législations. Michel Dubois a ainsi comparé le développement des recherches sur les cellules souches en France et en Angleterre en tenant compte des différences dans le traitement juridiques de ces questions (P. Brunet & Dubois, 2012). Aussi, Jarno Hoekman (2012) s'est interrogé sur la standardisation à l'échelle mondiale d'un type particulier d'écrits scientifiques : les essais cliniques. Ces travaux récents montrent qu'il y aurait peut-être à envisager l'intégration de la question juridique dans le champ des *science studies*.

Ces interfaces avec la géographie des migrations, la géographie politique, la géographie des nouvelles technologies d'information et de communications, la géographie de la santé, la géographie historique et plus généralement les études sociales des sciences, ne sont que des voies parmi d'autres dans l'étendue des possibles qu'offre la prise en compte d'une géographie des activités scientifiques disposant d'une certaine autonomie à l'égard de la géographie de l'enseignement supérieur et de l'innovation. La réussite dans l'obtention de cette autonomie est souhaitable, non seulement d'un point de vue scientifique, mais aussi dans la mesure où il est plus que jamais nécessaire d'envisager la science autrement qu'à travers ses relations à l'économie. Dans le débat public, les discours médiatiques et politiques, la recherche reste encore essentiellement perçue comme un levier pour la croissance économique et l'emploi. Aussi, les politiques publiques d'aménagement procèdent comme si la recherche était un secteur économique parmi d'autres, un secteur concurrentiel dont on pourrait mesurer les performances à l'aune des indicateurs bibliométriques. Beaucoup d'idées reçues sont prises pour des réalités, comme l'idée qu'avec la mondialisation, il y aurait un avantage à la concentration territoriale des chercheurs, alors qu'une fois confrontées aux observations empiriques, ces considérations ne trouvent aucune vérification : d'une part, la mondialisation existe dans ce domaine mais traduit une densification des collaborations à différents niveaux géographiques et l'intégration d'un nombre croissant de lieux dans le système ; d'autre part, le volume de production scientifique issu d'un lieu donné dépend linéairement du nombre de chercheurs qui y travaillent, sans qu'il y ait de plus-value liée à la concentration spatiale des chercheurs.

BIBLIOGRAPHIE

BIBLIOGRAPHIE GÉNÉRALE

Les références précédées d'un astérisque désignent des études issues de méta-analyses ou sources secondaires (en conformité avec les normes APA). Tous les articles dont on a trouvé l'identifiant électronique en ligne (**Digital Object Identifier**) à l'aide de la ressource documentaire **CrossRef** ou à partir du document cité lui-même sont suivis de la mention « en ligne ». Le lecteur de la version numérique de la thèse est invité à cliquer sur cette annotation s'il souhaite obtenir le **doi** de l'article et/ou accéder directement à son contenu.

- AALBERS, M. B., & ROSSI, U. (2007). Beyond the Anglo-American hegemony in human geography: a European perspective. *GeoJournal*, 67(2), 137-147. En ligne.
- ABBOTT, A. D. (2011). Time matters. Traduction de l'épilogue par Claire Lemercier et Carine Ollivier. *Terrains & Travaux*, 2(19), 183-203.
- ABEL, G. J., & SANDER, N. (2014). Quantifying Global International Migration Flows. *Science*, 343(6178), 1520-1522. En ligne.
- ACADÉMIE DES SCIENCES (FRANCE). (1922). Mémoires et communications des membres et des correspondants de l'Académie (Gallica). *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 175, 5-6.
- ACS, Z., AUDRETSCH, D. B., & FELDMAN, M. P. (1992). Real Effects of Academic Research: Comment. *The American Economic Review*, 82(1), 363-367.
- ADAMS, J., & PENDLEBURY, D. (2010). *Global research report: United States*. Leeds : Evidence. 8 p.
- ADAMS, J., & THOMSON, S. (2011). *Funding research excellence: research group size, critical mass & performance*. Londres : University Alliance. 40 p.
- ADDINK, E. A., VAN COILLIE, F. M. B., & DE JONG, S. M. (2012). Introduction to the GEOBIA 2010 special issue: From pixels to geographic objects in remote sensing image analysis. *International Journal of Applied Earth Observation and Geoinformation*, 15, 1-6. En ligne.
- ADELL-GOMBERT, N. (2008, juin 25). Une volonté de savoirs. *EspaceTemps.net*. En ligne.
- AGRAWAL, A., COCKBURN, I., & MCHALE, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, 6(5), 571-591. En ligne.
- AGRAWAL, A., KAPUR, D., MCHALE, J., & OETTL, A. (2011). Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, 69(1), 43-55. En ligne.
- ALLISON, P. D., & LONG, J. S. (1987). Interuniversity Mobility of Academic Scientists. *American Sociological Review*, 52(5), 643-652. En ligne.

- ALMEIDA, P., & KOGUT, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, 45(7), 905-917. En ligne.
- AMIN, A., & THRIFT, N. (1992). Neo-Marshallian Nodes in Global Networks*. *International journal of urban and regional research*, 16(4), 571-587. En ligne.
- ANDERSON, W. (2002). Introduction: Postcolonial Technoscience. *Social Studies of Science*, 32(5/6), 643-658. En ligne.
- ANDERSSON, Å. E., & PERSSON, O. (1993). Networking scientists. *The Annals of Regional Science*, 27(1), 11-21. En ligne.
- ANDERSSON, D. E., GUNESSEE, S., MATTHIESSEN, C. W., & FIND, S. (2014). The geography of Chinese science. *Environment and Planning A*, 46(12), 2950-2971. En ligne.
- ANDURAND, A., JÉGOU, L., MAISONOBE, M., & SIGRIST, R. (2014). Les mondes savants et leur visualisation, de l'Antiquité à aujourd'hui. *Histoire et Informatique*, à paraître.
- ANSELIN, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93-115. En ligne.
- ANSELIN, L., VARGA, A., & ACS, Z. (1997). Local Geographic Spillovers between University Research and High Technology Innovations. *Journal of Urban Economics*, 42(3), 422-448. En ligne.
- ANTHEAUME, B., DOLLFUS, O., DURAND-DASTÈS, F., BRUNET, R., GRASLAND, C., GRATALOUP, C., ... TAILLARD, C. (1999). La mondialisation. *Espace géographique*, 28(1), 36-48. En ligne.
- APPELT, S., VAN BEUZEKOM, B., GALINDO-RUEDA, F., & DE PINHO, R. (2015). *Which factors influence the international mobility of research scientists?* (Document de travail No. 2015/02). Paris : OCDE. 31 p. En ligne.
- ARCHAMBAULT, E. (2010). *30 years in Science. Secular Movements in Knowledge Creation*. Montréal : Science-Metrix. 16 p.
- ARRAULT, J.-B. (2007). *Penser à l'échelle du Monde. Histoire conceptuelle de la mondialisation en géographie (fin du XIXe siècle/entre-deux-guerres)* (Thèse de géographie sous la direction de Marie-Claire Robic). Université Panthéon-Sorbonne-Paris I, Paris, 706 p.
- ARTHUR, W. B. (1988). *Urban Systems and Historical Path-dependence* (Document de travail No. 0012). Stanford : Stanford Institute for Population and Resource Studies. 11 p.
- ARUNACHALAM, S. (2005). Science on the periphery: bridging the information divide. In H. F. MOED, W. GLÄNZEL, & U. SCHMOCH, *Handbook of Quantitative Science and Technology Research* (pp. 163-183). Dordrecht : Springer Netherlands. En ligne.
- ARVANITIS, R. (1983). *L'évaluation de la recherche aux États-Unis et l'utilisation de mesures quantitatives de la science*. (Rapport de recherche pour le Centre de Prospective et d'Évaluation, Ministère de l'industrie et de la recherche). Paris : Centre de Sociologie de l'Innovation. Ecole des Mines. 263 p.
- ARVANITIS, R. (2011). Que des réseaux ! Compte rendu de Caroline Wagner. The New Invisible College. Science for Development. *Revue d'anthropologie des connaissances*, 5(1), 178-185. En ligne.

- ARVANITIS, R., GROSSETTI, M., RAJ, K., RENAUD, P., & THOMAS, F. (2008). Sciences, savoirs et mondialisations. *Les Cahiers du M.U.R.S.*, (57/58), 51-69.
- ARVANITIS, R., WAAST, R., & GAILLARD, J. (2000). Science in Africa: A Bibliometric Panorama Using PASCAL Database. *Scientometrics*, 47(3), 457-473. En ligne.
- ASHEIM, B. T., & COENEN, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research Policy*, 34(8), 1173-1190. En ligne.
- ASTON, L., & SHUTT, L. (2009). *Concentration and Diversity : understanding the relationship between excellence, concentration and critical mass in UK research* (Research Paper No. 2009/01). Londres : University Alliance. 28 p.
- ATKINSON, P., BATCHELOR, C., & PARSONS, E. (1998). Trajectories of collaboration and competition in a medical discovery. *Science Technology and Human Values*, 23(3), 259-284. En ligne.
- AUBERT, F. D'. (2008). *Vers un partenariat renouvelé entre organismes de recherche, universités et grandes écoles*. (Rapport à Madame la ministre de l'Enseignement supérieur et de la Recherche). Paris : Ministère de l'Enseignement Supérieur et de la Recherche. 36 p.
- AUDRETSCH, D. B., & VIVARELLI, M. (1994). Small firms and R&D spillovers : Evidence from Italy. *Revue d'économie industrielle*, 67(1), 225-237. En ligne.
- AUST, J., & CRESPIY, C. (2009). Napoléon renversé ? Institutionnalisation des Pôles de recherche et d'enseignement supérieur et réforme du système académique français. *Revue française de science politique*, 59(5), 915. En ligne.
- AUTANT-BERNARD, C. (2001). Science and knowledge flows: evidence from the French case. *Research policy*, 30(7), 1069-1078. En ligne.
- AUTANT-BERNARD, C., BILLAND, P., & MASSARD, N. (2010). « L'économie industrielle depuis 30 ans : réalisations et perspectives ». *Innovation et espace - des externalités aux réseaux*. *Revue d'économie industrielle*, (129-130), 203-236. En ligne.
- AUTANT-BERNARD, C., MAIRESSE, J., & MASSARD, N. (2007). Spatial knowledge diffusion through collaborative networks. *Papers in Regional Science*, 86(3), 341-350. En ligne.
- AUTANT-BERNARD, C., & MASSARD, N. (2001). *Externalités de connaissances et géographie de l'innovation : les enseignements des études empiriques* (Document de travail CREUSET). Saint-Etienne : Université Jean Monnet Saint-Etienne. 24 p. En ligne.
- AVRAMIDES, A. (2006). Thomas Nagel: The view from nowhere. In J. SHAND (DIR.), *Central Works of Philosophy, Volume 5: The Twentieth Century: Quine and After* (Vol. 5, pp. 227-245). Montréal-Kingston : McGill-Queen's University Press.
- AYDALOT, P. (1965). Note sur les économies externes et quelques notions connexes. *Revue économique*, 16(6), 944. En ligne.
- BACON, F. (1627). *La nouvelle Atlantide*. (M. LE DOEUFF & M. LLASERA, TRAD.) (2^e éd.). Paris : Flammarion. 2000.
- BAHOKEN, F. (2011). *Représentation graphique des matrices. Graphe et/ou carte des flux?*. Paris : Groupe fmnr. 14 p. En ligne.

- BAHOKEN, F., BEAUGUITTE, L., & LHOMME, S. (2013a). *La visualisation des réseaux. Principes, enjeux et perspectives*. Paris : Groupe fnmr. 17 p. En ligne.
- BAHOKEN, F., BEAUGUITTE, L., & LHOMME, S. (2013b). Visualizing world flows: a challenge between efficacy, accuracy and aesthetics (3 p.). Présenté à 18e colloque européen de géographie quantitative et théorique (ECTQG 2013), Dourdan. En ligne.
- BAILLY, A. S. (1979). [Recension de *Idéologie, science et géographie humaine - Ideology, Science and human Geography*, par D. Gregory]. *Espace géographique*, 8(2), 104-104.
- BAILLY, A. S. (1985). Le rôle des activités de service dans le développement régional suisse. *Annales de géographie*, 94(523), 257-269. En ligne.
- BAILLY, A. S., DERYCKE, P.-H., & TORRE, A. (2012). *50 ans de science régionale francophone*. Paris : Economica pour l'Association de Science Régionale de Langue Française. 42 p.
- BAKIS, H. (1980). Éléments pour une géographie des télécommunications. *Annales de Géographie*, 89(496), 657-688. En ligne.
- BALLAND, P.-A. (2011). Proximity and the Evolution of Collaboration Networks: Evidence from Research and Development Projects within the Global Navigation Satellite System (GNSS) Industry. *Regional Studies*, 46(6), 741-756. En ligne.
- BALLAND, P.-A., BOSCHMA, R. A., & FRENKEN, K. (2013). *Proximity and Innovation: From Statics to Dynamics* (Papers in Evolutionary Economic Geography No. 13.14). Utrecht : Utrecht University. 26 p. En ligne.
- BARJAVEL. (1968). *La nuit des temps*. Paris : Pocket. 2012. 410 p.
- BARNES, T. J. (2001). Rethorizing economic geography: from the quantitative revolution to the « cultural turn ». *Annals of the Association of American Geographers*, 91(3), 546-565. En ligne.
- BARO, J., MERING, C., & VACHIER, C. (2014). Peut-on cartographier des taches urbaines à partir d'images Google Earth ? Une expérience réalisée à partir d'images de villes d'Afrique de l'Ouest. *Cybergeog: European Journal of Geography*, (682). En ligne.
- BARON, M. (1994). *La place des IUT dans la structuration de l'espace de formation supérieure en France* (Thèse de géographie sous la direction de Thérèse Saint-Julien). Université Paris I-Panthéon-Sorbonne, Paris.
- BARON, M. (2010). Les transformations de la carte universitaire depuis les années 1960 : constats et enjeux. *Le Mouvement Social*, 233(4), 93-105. En ligne.
- BARON, M. (2012). *Mises en espace des sociétés de la connaissance par les universités et les mobilités étudiantes : position et projet scientifique* (Habilitation à diriger des recherches). Université Paris 7-Paris Diderot, Paris, 300 p.
- BARON, M., BARRERA, C., & BIRCK, F. (2015). Universités et Territoires, du passé faisons table rase? *Les Annales de la recherche urbaine*, (109).
- BARON, M., & BERROIR, S. (2007). Paris et le système universitaire français : mythe et réalités. *Annales de Géographie*, (655), 3-22. En ligne.

- BARON, M., CUNNINGHAM-SABOT, E., GRASLAND, C., RIVIÈRE, D., & VAN HAMME, G. (2010). *Villes et régions européennes en décroissance. Maintenir la cohésion territoriale*. Paris : Hermès. 346 p.
- BARON, M., ECKERT, D., & JÉGOU, L. (2011). Peut-on démêler l'écheveau mondial des collaborations scientifiques? *M@ppemonde*, 2(102). En ligne.
- BARON, M., ECKERT, D., & MAISONOBE, M. (2011). Worlds of scientific cooperation. Choosing appropriate data for specific networks. (9 p.). Présenté à 17e colloque européen de géographie quantitative et théorique (ECTQG 2011), Athènes. En ligne.
- BARON, M., LE NEINDRE, C., & ZANIN, C. (2008). Mettre en cartes les universités du Bassin parisien. *M@ppemonde*, 4(92). En ligne.
- BARON, M., & PERRET, C. (2006). Bacheliers, étudiants et jeunes diplômés: quels systèmes migratoires régionaux? *L'Espace géographique*, (1), 44-62. En ligne.
- BARRÉ, R., HERNÁNDEZ, V., MEYER, J.-B., & VINCK, D. (2003). *Diasporas scientifiques/Scientific Diasporas*. Paris : Presses de l'IRD. 197 p.
- BARUFFALDI, S. H., & LANDONI, P. (2012). Return mobility and scientific productivity of researchers working abroad: The role of home country linkages. *Research Policy*, 41(9), 1655-1665. En ligne.
- BASALLA, G. (1967). The Spread of Western Science. *Science*, 156(3775), 611-622. En ligne.
- BASSECOULARD, E., & ZITT, M. (1998). Internationalization of scientific journals: A measurement based on publication and citation scope. *Scientometrics*, 41(1-2), 255-271. En ligne.
- BASTIÉ, J., & BRICHLER, M. (1960). Délimitation de l'agglomération parisienne. *Population*, 15(3), 433-456. En ligne.
- BASU, A. (2010). Does a country's scientific 'productivity' depend critically on the number of country journals indexed? *Scientometrics*, 82(3), 507-516. En ligne.
- BATAGELJ, V., KEJZAR, N., & KORENJAK-CERNE, S. (2006). Analyzing the Structure of U.S. Patents Network. In *Data Science and Classification* (140-148 p.). Ljubljana : Springer.
- BATHELT, H. (2007). Buzz-and-Pipeline Dynamics: Towards a Knowledge-Based Multiplier Model of Clusters. *Geography Compass*, 1(6), 1282-1298. En ligne.
- BATHELT, H., MALMBERG, A., & MASKELL, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31-56. En ligne.
- BEAUGUITTE, L. (2011, octobre 20). *L'Assemblée générale de l'ONU de 1985 à nos jours : acteur et reflet du Système-monde. Essai de géographie politique quantitative*. Université Paris Diderot 7, Paris, 313 p.
- BEAUGUITTE, L. (2012). *Analyser les réseaux avec R (packages statnet, igraph et tnet)*. Paris : Groupe fm̄r. 17 p. En ligne.

- BEAUGUITTE, L., & BEAUGUITTE, P. (2011). *Opérations matricielles et analyse de graphe*. Paris : Groupe fnmr. 11 p. En ligne.
- BEAUGUITTE, L., GUÉRIN-PACE, F., & RICHARD, Y. (2014). L'UE et ses voisinages : acteurs et espaces d'un discours. In *Fronts et frontières des sciences du territoire* (2^e colloque international du CIST, pp. 25-31). Paris : Collège international des sciences du territoire.
- BEAUJEU-GARNIER, J. (1980). *Géographie urbaine* (2^e éd.). Paris : Armand Colin. 1983. 360 p.
- BEAVER, D., & ROSEN, R. (1978). Studies in scientific collaboration: Part I. The professional origins of scientific co-authorship. *Scientometrics*, 1(1), 65-84. En ligne.
- BECKMANN, M. (1993). Knowledge networks: the case of scientific interaction at a distance. *The Annals of Regional Science*, 27(1), 5-9. En ligne.
- BELL, D., & JAYNE, M. (2009). Small cities? Towards a research agenda. *International Journal of Urban and Regional Research*, 33(3), 683-699. En ligne.
- BELLÙ, L. G., & LIBERATI, P. (2006). *Analyse d'inégalité. L'indice de gini* (Outils Analytiques No. Easypol 040). ONU. 30 p. En ligne.
- BENAIM, M. (2013, mai 3). *Gouvernance et politiques de science en région : une approche multi-acteurs et multi-niveaux* (Thèse de Sciences Economiques sous la direction de Jean-Alain Héraud). Université de Strasbourg, Strasbourg, 456 p.
- BEN-DAVID, J. (1970). Sociologie de la science : introduction. *Revue internationale des sciences sociales*, 22(1), 7-29.
- BEN-DAVID, J. (1991). *Scientific growth: essays on the social organization and ethos of science*. (G. FREUDENTHAL, ÉD.). Los Angeles : Univ. of California Press.
- BEN-DAVID, J., & COLLINS, R. (1966). Social Factors in the Origins of a New Science: The Case of Psychology. *American Sociological Review*, 31(4), 451-465. En ligne.
- BENKO, G. (1990). La cité scientifique d'Ile-de-France-Sud. *Annales de Géographie*, 99(556), 660-674. En ligne.
- BENKO, G. (2007). Économie urbaine et régionale au tournant du siècle. *Métropoles*, (1), 141-181.
- BENKO, G., & LIPIETZ, A. (1992). *Les régions qui gagnent : districts et réseaux : les nouveaux paradigmes de la géographie économique*. Paris : Presses universitaires de France. 424 p.
- BERNELA, B., BOUBA-OLGA, O., & FERRU, M. (2013). *Geographical mobility of PhDs: an analysis of French long-term data* (Document de travail). Poitiers : CRIEF Université de Poitiers. 18 p. En ligne.
- BERRY, B. J. L. (1964). Approaches to Regional Analysis: A Synthesis. *Annals of the Association of American Geographers*, 54(1), 2-11. En ligne.
- BERRY, B. J. L. (1973). A paradigm for Modern Geography. In *Directions in geography* (pp. 3-21). Londres : Methuen.

- BERRY, B. J. L., GOHEEN, P. G., & GOLDSTEIN, H. (1969). *Metropolitan area definition: a re-evaluation of concept and statistical practice* (No. 28). Washington, D.C. : U.S. Bureau of the Census. 45 p.
- BÈS, M.-P., & GROSSETTI, M. (2003). Dynamiques des réseaux et des cercles. Encastres et découplages. *Revue d'économie industrielle*, 43-58.
- BESSE, J.-M. (2010). Approches spatiales dans l'histoire des sciences et des arts. *L'Espace géographique*, (3), 211-224.
- BIDART, C. (2008). Dynamiques des réseaux personnels et processus de socialisation : évolutions et influences des entourages lors des transitions vers la vie adulte. *Revue française de sociologie*, 49(3), 559-583. En ligne.
- BIDART, C., & DUPRAY, A. (2014). Life Course, Time and Process: Exploring « Trans- » Quantitative and Qualitative Methods. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 124(1), 5-13. En ligne.
- BIDART, C., & DUPRAY, A. (2015). Narratives and Traces: « Trans- » Quantitative and Qualitative Methods for Bridging the Gap between Objective and Subjective Data. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 125(1), 5-7. En ligne.
- BITBOL, M. (1998). Le réel-en-soi, l'inconnaissable, et l'ineffable. *Annales d'Histoire et de Philosophie du Vivant*, (1), 143-152.
- BLICKENSTAFF, J., & MORAVCSIK, M. J. (1982). Scientific output in the third world. *Scientometrics*, 4(2), 135-169. En ligne.
- BLONDEL, V., GUILLAUME, J.-L., LAMBIOTTE, R., & LEFEBVRE, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, 2008(10), p. 10008. En ligne.
- BLONDEL, V., KRINGS, G., & THOMAS, I. (2010). Régions et frontières de téléphonie mobile en Belgique et dans l'aire métropolitaine bruxelloise. *Brussels studies*, (42).
- BLOOR, D. (1976). *Knowledge and social imagery*. Londres-Boston : Routledge & K. Paul. 156 p.
- BOMSEL, O. (2010). *L'économie immatérielle : Industries et marchés d'expériences*. Paris : Gallimard. 288 p.
- BONACCORSI, A., & DARAIO, C. (2005). Exploring size and agglomeration effects on public research productivity. *Scientometrics*, 63(1), 87-120. En ligne.
- BOOTSMA, D. (2001). The Dutch DNA Repair Group in Retrospect. *Mutation Research*, 485(1), 37-41. En ligne.
- BÖRNER, K., PENUMARTHY, S., MEISS, M., & KE, W. (2006). Mapping the diffusion of scholarly knowledge among major U.S. research institutions. *Scientometrics*, 68(3), 415-426. En ligne.
- BÖRNER, K., SANYAL, S., & VESPIGNANI, A. (2007). Network science. *Annual Review of Information Science and Technology*, 41(1), 537-607. En ligne.

- BORNMAN, L., & LEYDESDORFF, L. (2012). Which are the best performing regions in information science in terms of highly cited papers? Some improvements of our previous mapping approaches. *Journal of Informetrics*, 6(2), 336-345. En ligne.
- BORNMAN, L., LEYDESDORFF, L., WALCH-SOLIMENA, C., & ETTL, C. (2011). Mapping excellence in the geography of science: An approach based on Scopus data. *Journal of Informetrics*, 5(4), 537-546. En ligne.
- BORNMAN, L., & PLUME, A. (2011). Is it necessary to consider suburbs (or small cities in the close proximity) and name variants in a citation impact analysis for bigger cities? An investigation using Munich as an example. *Journal of Informetrics*, 5(4), 695-697. En ligne.
- BORNMAN, L., & WALTMAN, L. (2011). The detection of « hot regions » in the geography of science—A visualization approach by using density maps. *Journal of Informetrics*, 5(4), 547-553. En ligne.
- BOSCHMA, R. A. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1), 61-74. En ligne.
- BOSCHMA, R. A., & FRENKEN, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, 6(3), 273-302. En ligne.
- BOUBA-OLGA, O., & GROSSETTI, M. (2008). Socio-économie de proximité. *Revue d'économie régionale & urbaine*, (3), 311-328. En ligne.
- BOUBA-OLGA, O., & GROSSETTI, M. (2012). L'excellence des Universités : oui, mais quelle excellence ?! *Cahiers Français*, (138), 38-39.
- BOUDON, R. (2008). *Le relativisme*. Paris : Presses universitaires de France. 127 p.
- BOURDIEU, P. (1976). Le champ scientifique. *Actes de la recherche en sciences sociales*, 2(2), 88-104. En ligne.
- *BOURDIEU, P. (2000). *Les structures sociales de l'économie*. Paris : Seuil. 296 p.
- BOURDIEU, P. (2002a). Les conditions sociales de la circulation internationale des idées. *Actes de la recherche en sciences sociales*, 145(5), 3-8. En ligne.
- BOURDIEU, P. (2002b). *Science de la science et réflexivité cours du Collège de France (2000-2001)*. Paris : Raisons d'agir. 240 p.
- BOURDIN, S. (2014). Les défis de la future politique régionale européenne 2014-2020 : on prend les mêmes et on recommence ? *Géoconfluences*. En ligne.
- BOYACK, K. W., KLAIVANS, R., & BÖRNER, K. (2005). Mapping the backbone of science. *Scientometrics*, 64(3), 351-374. En ligne.
- BRAUDEL, F. (1985). *La dynamique du capitalisme*. Paris : Flammarion. 2008. 128 p.
- BRAUNSTEIN, J.-F. (2003). Thomas Kuhn lecteur de Ludwik Fleck. *Archives de Philosophie*, (3), 403-422.
- BRAUN, T., GLÄNZEL, W., & SCHUBERT, A. (2010). On Sleeping Beauties, Princes and other tales of citation distributions.... *Research Evaluation*, 19(3), 195-202. En ligne.

- BRESCHI, S., & LISSONI, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, 9(4), 439-468. En ligne.
- BRETAGNOLLE, A., DELISLE, F., MATHIAN, H., LIZZI, L., GUÉROIS, M., & AVERLANT, G. (2011). *LUZ specifications (Urban Audit 2004)* (Technical Report, Espon 2013 Database). Paris : Union Européenne. 74 p. En ligne.
- BRETAGNOLLE, A., & PUMAIN, D. (2010). Simulating Urban Networks through Multiscalar Space-Time Dynamics: Europe and the United States, 17th-20th Centuries. *Urban Studies*, 47(13), 2819-2839. En ligne.
- BROCARD, M. (1978). *Recherche scientifique et développement régional en France*. (Thèse d'Etat sous la direction de Jacqueline Beaujeu-Garnier). Paris, 628 p.
- BROCARD, M. (1981). Aménagement du territoire et développement régional : le cas de la recherche scientifique. *Espace géographique*, 10(1), 61-73. En ligne.
- BROCARD, M. (1986). La recherche scientifique, enjeu du développement régional ? [Table ronde organisée par le CNRS à Montpellier le 13 janvier 1986]: Table ronde organisée par le CNRS à Montpellier le 13 janvier 1986. *Espace géographique*, 15(4), 302-306. En ligne.
- BROCARD, M. (1991). *La science et les régions : géoscopie de la France*. Montpellier-Paris : RECLUS-La Documentation Française. 267 p.
- BROCARD, M., HÉRIN, R., & JOLY, J. (1996). *Atlas de France. Volume 4, Formation et Recherche*. Montpellier-Paris : RECLUS-La Documentation Française. 128 p.
- BROEKEL, T., & BOSCHMA, R. A. (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, 12(2), 409-433. En ligne.
- BRUNET, P., & DUBOIS, M. (2012). Cellules souches et technoscience : sociologie de l'émergence et de la régulation d'un domaine de recherche biomédicale en France. *Revue française de sociologie*, 53(3), 391-428.
- BRUNET, R. (1986). La géographie dite « sociale » : fonctions et valeurs de la distinction. *Espace géographique*, 127-130. En ligne.
- BRUNET, R. (1990). Point de vue : l'université, la ville et la région. *Espace géographique*, 213-213. En ligne.
- BRUNET, R. (1992). *Les Mots de la géographie : dictionnaire critique*. Montpellier-Paris : RECLUS-La Documentation Française. 472 p.
- BRUNET, R. (2001). *Le déchiffrement du monde : théorie et pratique de la géographie*. Paris : Belin. 400 p.
- BRUNET, R. (2003). *Le diamant : un monde en révolution*. Paris : Belin. 416 p.
- BRUNET, R., & DOLLFUS, O. (1990). *Géographie Universelle : Mondes Nouveaux*. Montpellier-Paris : RECLUS-Belin. 552 p.

- BRUTEL, C., & LEVY, D. (2011). *Le nouveau zonage en aires urbaines de 2010. 95 % de la population vit sous l'influence des villes* (INSEE Première No. 1374). Paris : INSEE. 4 p. En ligne.
- BUELA-CASAL, G., PERAKAKIS, P., TAYLOR, M., & CHECA, P. (2006). Measuring internationality: Reflections and perspectives on academic journals. *Scientometrics*, 67(1), 45-65. En ligne.
- BUSH, V. (1945, juin). As We May Think. *The Atlantic Online*. En ligne.
- BUSH, V. (1960). *Science, the endless frontier: a report to the President on a program for postwar scientific research*. Washington, D.C. : National Science Foundation. 193 p.
- CAILLÉ, A., & DUFOIX, S. (2013). *Le tournant global des sciences sociales*. Paris : La Découverte. 441 p.
- CALAS, B. (2013, janvier). *Le monde de la rose, une filière mondiale*. Café géographique présenté à Les Cafés géographiques à Toulouse, Papagayo. En ligne.
- CALBÉRAC, Y. (2010, décembre 13). *Terrains de géographes, géographes de terrain. Communauté et imaginaire disciplinaires au miroir des pratiques de terrain des géographes français du XXe siècle*. (Thèse de géographie sous la direction d'Isabelle Lefort). Université Lumière-Lyon II, Lyon, 390 p.
- CALLON, M., & FORAY, D. (1997). Introduction : Nouvelle économie de la Science ou socio-économie de la recherche scientifique? *Revue d'économie industrielle*, 79(1), 13-35.
- CAMAGNI, R., & CAPELLO, R. (2010). Macroeconomic and territorial policies for regional competitiveness: an EU perspective. *Regional Science Policy & Practice*, 2(1), 1-19. En ligne.
- CAMPBELL, D., LEFEBVRE, C., PICARD-AITKEN, M., CÔTÉ, G., VENTIMIGLIA, A., ROBERGE, G., & ARCHAMBAULT, É. (2013). *Country and Regional Scientific Production Profiles* (Rapport de Science-Metrix pour la Commission Européenne). Luxembourg : Publications Office of the European Union. 200 p.
- CAMPBELL, D., VENTIMIGLIA, A., & ARCHAMBAULT, É. (2013). *Scientific output and collaboration of companies publishing the most in the ERA* (Rapport de Science-Metrix pour la Commission Européenne). Luxembourg : Publications Office of the European Union. 44 p. En ligne.
- CAÑIBANO, C., OTAMENDI, F. J., & SOLÍS, F. (2011). International temporary mobility of researchers: a cross-discipline study. *Scientometrics*, 89(2), 653-675. En ligne.
- CARO, P. (1994). *Le rôle de la formation, de la recherche et de l'innovation industrielle dans l'organisation de l'espace du Centre-Est français* (Thèse de géographie sous la direction de Robert Chapuis). Université de Dijon.
- CARO, P. (2003). L'apport de la géographie à la relation formation-emploi. *Formation Emploi*, (84).
- CARO, P. (2006). La dimension spatiale des systèmes formation-emploi. *L'Espace géographique*, (3), 223-240.

- CARPENTER, M. P., & NARIN, F. (1981). The adequacy of the Science Citation Index (SCI) as an indicator of international scientific activity. *Journal of the American Society for Information Science*, 32(6), 430-439. En ligne.
- CARRÉ, M.-N. (2013). Le système global migratoire : une énigme polysémique? *M@ppemonde*, 1(112). En ligne.
- CARRINCAZEAX, C., LUNG, Y., & VICENTE, J. (2008). The Scientific Trajectory of the French School of Proximity: Interaction- and Institution-based Approaches to Regional Innovation Systems. *European Planning Studies*, 16(5), 617-628. En ligne.
- CARROUÉ, L. (1993). Berlin réunifiée : une nouvelle métropole à vocation internationale en Europe centrale. *Annales de Géographie*, 102(570), 113-130. En ligne.
- CARROUÉ, L. (2007). *Géographie de la mondialisation* (3^e éd.). Paris : Armand Colin. 256 p.
- CASTELLS, M. (1996). *The Rise of The Network Society: The Information Age: Economy, Society and Culture* (2^e éd.). Oxford : Wiley Blackwell. 2000. 594 p.
- CATTAN, N. (2007). *Cities and Networks in Europe: A critical approach of polycentrism*. Montrouge : John Libbey. 224 p.
- CATTAN, N., PUMAIN, D., ROZENBLAT, C., & SAINT-JULIEN, T. (1999). *Le système des villes européennes*. Paris : Economica/Anthropos. 197 p.
- CATTAN, N., & ROZENBLAT, C. (1991). Délimiter et mesurer l'évolution des populations urbaines. *Population*, 46(4), 979-991.
- CENTER FOR INTERNATIONAL EARTH SCIENCE INFORMATION NETWORK - CIESIN - COLUMBIA UNIVERSITY, UNITED NATIONS FOOD AND AGRICULTURE PROGRAMME - FAO, et CENTRO INTERNACIONAL DE AGRICULTURA TROPICAL - CIAT. (2005). Gridded Population of the World, Version 3 (GPWv3): Population Count Grid. NASA Socioeconomic Data and Applications Center (SEDAC). En ligne.
- CHABAULT, D. (2006). Les systèmes territoriaux de production : revue de littérature et approches théoriques d'un concept évolutif. *Cahiers de Recherche du CERMAT*, 19(6).
- CHALAYE, S., & MASSARD, N. (2012). *Géographie de l'innovation en Europe : observer la diversité des régions françaises*. Paris : la Documentation française. 112 p.
- CHAMBERS, D. W., & GILLESPIE, R. (2000). Locality in the History of Science: Colonial Science, Technoscience, and Indigenous Knowledge. *Osiris*, 15, 221-240.
- CHARLES, C. (Dir.). (2009). *Le temps des capitales culturelles XVIIIe-XXe siècles*. Seyssel (Ain) : Champ Vallon. 400 p.
- CHARRE, J. (1995). *Statistique et Territoire*. Montpellier : RECLUS. 120 p.
- CHASE, I. D. (1982). Dynamics of hierarchy formation: the sequential development of dominance relationships. *Behaviour*, 80(3-4), 3-4. En ligne.
- CHESHIRE, P. (1997). Economic indicators for European cities and regions: why boundaries matter (15 p.). Présenté à Eurostat Seminar on Urban Indicators, Barcelone : 14-16 Avril 1997. En ligne.

- CHEVALIER, M. (Dir.). (1989). La Géographie de la créativité et de l'innovation : actes du colloque tenu les 9-10 juin 1987 à l'Université de Paris-Sorbonne. Paris : Université de Paris-Sorbonne, Dép. de géographie.
- CHINCHILLA-RODRIGUEZ, Z., VARGAS-QUESADA, B., HASSAN-MONTERO, Y., GONZALEZ-MOLINA, A., & MOYA-ANEGON, F. (2010). New approach to the visualization of international scientific collaboration. *Information Visualization*, 9(4), 277-287. En ligne.
- CHOMSKY, N. (2009, février 9). Chomsky: Understanding the Crisis Markets, the State and Hypocrisy [Foreign Policy in Focus]. En ligne.
- CHUBIN, D. E. (1985). Beyond invisible colleges: Inspirations and aspirations of post-1972 social studies of science. *Scientometrics*, 7(3), 221-254. En ligne.
- CLAVAL, P. (1972). La réflexion théorique en géographie et les méthodes d'analyse. *Espace géographique*, 7-22. En ligne.
- CLAVAL, P. (1977). *La nouvelle géographie* (1^{re} éd.). Paris : Presses Universitaires de France. 128 p.
- CLAVAL, P. (1988). [Recension de *À propos de Denys d'Alexandrie - Géographie et culture en Grèce ancienne. Essai de lecture de la « description de la terre habitée » de Denys d'Alexandrie. Paris, Thèse de doctorat d'État, École pratique des hautes études en sciences sociales*, par C. Jacob]. *Espace géographique*, 17(1), 53-54.
- CLEAVER, J. E., KARPLUS, K., KASHANI-SABET, M., & LIMOLI, C. L. (2001). Nucleotide excision repair « a legacy of creativity ». *Mutation Research*, 485, 23-36. En ligne.
- CLERC, P. (2007). Émile Levasseur, un libéral en géographie. *L'Espace géographique*, (1), 79-92.
- COHEN-TANUGI, L. (2009, décembre 25). L'avenir de la stratégie de Lisbonne : l'Europe ne doit pas abdiquer ses ambitions. *Le Monde*. En ligne.
- COMIN, M.-N. (2009). *Réseaux de villes et réseaux d'innovation en Europe: structuration du système des villes européennes par les réseaux de recherche sur les technologies convergentes* (Thèse de géographie sous la direction de Denise Pumain). Université Paris I-Panthéon-Sorbonne, Paris, 724 p.
- COMIN, M.-N. (2011). Les villes européennes dans les réseaux de recherche de pointe. In D. PUMAIN & M.-F. MATTEI, *Données Urbaines* (Vol. 6, pp. 289-295). Paris : Economica/Anthropos.
- COOK, H. J. (2008). Amsterdam, entrepôt des savoirs au XVII^e siècle. *Revue d'histoire moderne et contemporaine*, (2), 19-42.
- CRANE, D. (1969a). La diffusion des innovations scientifiques. *Revue française de sociologie*, 166-185. En ligne.
- CRANE, D. (1969b). Social Structure in a Group of Scientists: A Test of the « Invisible College » Hypothesis. *American Sociological Review*, 34(3), 335-352. En ligne.
- CRANE, D. (1972). *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago : University of Chicago Press. 213 p.

- CRAWFORD, E., SHINN, T., & SÖRLIN, S. (1993). The Nationalization and Denationalization of the Sciences: An Introductory Essay. In E. CRAWFORD, T. SHINN, & S. SÖRLIN, *Denationalizing Science. The Contexts of International Scientific Practice* (pp. 1-42). Dordrecht : Springer Netherlands.
- CREAGER, A. N. H. (1999). Book Review: Errol C. Friedberg. Correcting the Blueprint of Life: An Historical Account of the Discovery of DNA Repair Mechanisms. *Bulletin of the History of Medicine*, 3(73), 541-543.
- CREAGER, A. N. H. (2002). Tracing the politics of changing postwar research practices: the export of 'American' radioisotopes to European biologists. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 33(3), 367-388. En ligne.
- CREAGER, A. N. H., & SANTESMASES, M. J. (2006). Radiobiology in the Atomic Age: Changing Research Practices and Policies in Comparative Perspective. *Journal of the History of Biology*, 39(4), 637-647. En ligne.
- CRESPY, C., & SOLDANO, C. (2015). Institutions et politiques universitaires : quelle(s) intégration(s) des sites «secondaires»? In R. LEVY, C. SOLDANO, & P. CUNTIGH, *L'université et ses territoires. Dynamismes des villes moyennes et particularités de sites*. (pp. 33-54). Grenoble : Presses universitaires de Grenoble.
- CUYALA, S. (2013). La diffusion de la géographie théorique et quantitative européenne francophone d'après les réseaux de communications aux colloques européens (1978-2011). *Cybergeo: European Journal of Geography*, (657). En ligne.
- CUYALA, S. (2014, octobre 20). *Analyse spatio-temporelle d'un mouvement scientifique. L'exemple de la géographie théorique et quantitative européenne francophone*. (Thèse de géographie sous la direction de Denise Pumain). Université Paris I-Panthéon-Sorbonne, Paris, 460 p.
- DARRÉON, J.-L., & FILÂTRE, D. (2003). *Les universités à l'heure de la gouvernance*. Presses universitaires du Mirail. 224 p.
- DASGUPTA, P., & DAVID, P. A. (1994). Toward a new economics of science. *Research Policy. Special Issue in Honor of Nathan Rosenberg*, 23(5), 487-521. En ligne.
- DASSA, M., DENIAU, A., DURAND-BARTHEZ, M., GIRARD, F., POTHIER, N., & SÉNÉ, A. (2014). Le millefeuille des affiliations françaises dans les publications scientifiques. *Documentaliste - Sciences de l'information*, 51(4), 12-16. En ligne.
- DAVEZIES, L. (2008). *La République et ses territoires. La circulation invisible des richesses*. (1^{re} éd.). Paris : Seuil. 110 p.
- DAVIET, S. (2005). Trente ans de géographie industrielle dans les Annales de géographie (1970-1999) // Thirty years of industrial geography in the Annales de géographie, 1970-1999. *Annales de Géographie*, 114(641), 73-92. En ligne.
- DE BERNY-RICHE, C. (2013). *Parcours et expériences d'études des Franciliens selon leur origine, Une exploitation de l'enquête « trajectoire et origines »* (No. 2.10.019). Paris : IAU-IdF. 27 p. En ligne.

- DE BERNY-RICHE, C., & MAISONOBE, M. (2009). *La mobilité internationale des étudiants français en 2006* (Note rapide IAU-IdF No. 494). Paris : IAU-IdF. 4 p. En ligne.
- DEBORD, G. (1967). *La Société du Spectacle*. Paris : Gallimard. 2011. 224 p.
- DEDIJER, S. (1966). The science of science: A programme and a plea. *Minerva*, 4(4), 489-504. En ligne.
- DÉJOIE-LARNAUDIE, A. (2011). *Les zonages en unités et aires urbaines 2010 : des pôles très denses, des couronnes d'influence très rurales* (INSEE flash Champagne-Ardenne No. 138). Reims : INSEE. 8 p. En ligne.
- DE MEO, P., FERRARA, E., FIUMARA, G., & PROVETTI, A. (2011). Generalized Louvain method for community detection in large networks. *under review*.
- DEVORET, R. (2001). At the Birth of Molecular Radiation Biology. *Environmental and Molecular Mutagenesis*, 38(2-3), 135-143. En ligne.
- DIDELON, C. (2010). Une course vers le ciel. Mondialisation et diffusion spatio-temporelle des gratte-ciel. *M@ppemonde*, 3(99). En ligne.
- DI MÉO, G. (2008). La géographie culturelle : quelle approche sociale ? *Annales de géographie*, 660-661(2), 47. En ligne.
- DOLLFUS, O. (1997). *La Mondialisation* (3^e éd.). Paris : Presses de Sciences Po. 2007. 166 p.
- DOREIAN, P. (1979). On the Evolution of Group and Network Structure. *Social Networks*, 2(3), 235-252. En ligne.
- DUBET, F., FILÂTRE, D., MERRIEN, F.-X., SAUVAGE, A., & VINCE, A. (1994). *Universités et villes*. Paris : L'Harmattan. 318 p.
- DUBOIS, M. (2001). *La nouvelle sociologie des sciences*. Paris : Presses Universitaires de France. 256 p.
- DUCRUET, C. (2004). *Les villes-ports : laboratoire de la mondialisation* (Thèse de géographie sous la direction de Madeleine Brocard). Université, Le Havre, 434 p.
- DUCRUET, C. (2010). *Les mesures locales d'un réseau*. Groupe fmr. 11 p. En ligne.
- DUPUY, G. (2007). *La fracture numérique*. Paris : Ellipses. 158 p.
- DUPUY, G. (2008). [Recension de *Cities and networks in Europe, A critical approach of polycentrism*, Montrouge, John Libbey, 207 p., par N. Cattani]. *Cybergeog: European Journal of Geography*. En ligne.
- DUSSOUY, G. (2011). Pragmatisme et géopolitique - Les opportunités méthodologiques d'une retrouvaille épistémologique. *L'Espace Politique*, (12). En ligne.
- ECKERT, D., & BARON, M. (2013). Construire une géographie de la science. *M@ppemonde*, 2(110). En ligne.
- ECKERT, D., BARON, M., & JÉGOU, L. (2013). Les villes et la science : apports de la spatialisation des données bibliométriques mondiales. *M@ppemonde*, 110(2). En ligne.
- ECKERT, D., GOUZÉVITCH, D., GOUZÉVITCH, I., & PANE, M.-N. (à paraître). *La Russie : construction et crise d'un système scientifique*. Chapitre à paraître.
- ECKERT, D., GROSSETTI, M., JÉGOU, L., & MAISONOBE, M. (à paraître). *Les villes de la science contemporaine, entre logiques locales, nationales et globales : une approche bibliométrique*. Chapitre à paraître.

- EDGE, D. O., & MULKAY, M. (1976). *Astronomy transformed: the emergence of radio astronomy in Britain*. New York : Wiley. 482 p.
- EISENHOWER, D. D. (2003). Atoms for peace. *IAEA BULLETIN*, 45(2), 62-67.
- ÉLOIRE, F. (2010). Le classement de Shanghai. Histoire, analyse et critique. *L'Homme et la société*, 178(4), 17. En ligne.
- ELZINGA, A. (2001). Science and Technology: Internationalization. In N. J. SMELSER & P. B. BALTES (DIR.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 13633-13638). Amsterdam : Elsevier.
- ELZINGA, A. (2012). Features of the current science policy regime: Viewed in historical perspective. *Science and Public Policy*, 39(4), 416-428. En ligne.
- EMANGARD, P.-H., & BROCARD, M. (1976). Recherche scientifique et évolution économique régionale : le cas de la Bretagne. *Cahiers géographiques de Rouen*, (5), 1-63.
- ESTADES, J., JOLY, P. B., & MANGEMATIN, V. (1996). Dynamique des relations industrielles dans les laboratoires d'un grand organisme public de recherche : Coordination, apprentissage, réputation et confiance. *Sociologie du travail*, 3, 391-407.
- EUROSTAT. (2010). *European Regional and Urban Statistics Reference Guide*. Luxembourg : Publications Office of the European Union. 263 p.
- FACHE, J. (1999). La définition des industries de haute technologie. *Méditerranée*, 92(3), 41-48. En ligne.
- FELDMAN, M. P. (1994). *The Geography of Innovation*. Dordrecht : Springer Netherlands. 159 p.
- FELOUZIS, G. (2001). Les délocalisations universitaires et la démocratisation de l'enseignement supérieur. *Revue française de pédagogie*, 136(1), 53-63. En ligne.
- FEN-CHONG, J., & PFAENDER, F. (2011, septembre). *Analysing foreigners mobility in Paris from mobile phone data*. Présenté à 17e colloque européen de géographie quantitative et théorique (ECTQG 2011), Athènes.
- FERLIE, E., MUSSELIN, C., & ANDRESANI, G. (2009). The Governance of Higher Education Systems: A Public Management Perspective. In C. PARADEISE, E. REALE, I. BLEIKLIE, & E. FERLIE (DIR.), *University Governance* (Vol. 25, pp. 1-19). Dordrecht : Springer Netherlands. En ligne.
- FERRU, M. (2009, novembre 18). *La géographie des collaborations pour l'innovation : Le rôle des contraintes de ressources et de mise en relation* (Thèse de Sciences Economiques sous la direction d'Olivier Bouba-Olga). Université de Poitiers, Poitiers, 351 p.
- FILÂTRE, D. (1994). La fin de l'hégémonie toulousaine ou l'apparition d'un nouveau système local d'enseignement supérieur. In M. GROSSETTI, *Université et Territoire. Un système local d'enseignement supérieur: Toulouse et Midi-Pyrénées* (pp. 181-212). Toulouse : Presses universitaires du Mirail.
- FILÂTRE, D. (2015). Postface. Reconnaître les formes territoriales de l'université. In *L'université et ses territoires. Dynamismes des villes moyennes et particularités de sites*. (pp. 151-163). Grenoble : Presses universitaires de Grenoble.

- FILÂTRE, D., MILARD, B., & VIÉS, G. (1996). *Université 2000, les effets des implantations nouvelles . Volume 1 . Etude sur la région Midi-Pyrénées*. Toulouse : DATAR, Certop. 319 p.
- *FLECK, L. (1935). *Genèse et développement d'un fait scientifique*. (N. JAS, TRAD.). Paris : Belles lettres. 2005. 280 p.
- FLORIDA, R. L. (2005a). *Cities and the creative class*. New York : Routledge. 193 p.
- FLORIDA, R. L. (2005b). The world is spiky. *The Atlantic*, 296(3), 48-51.
- FORAY, D. (2000). *L'économie de la connaissance*. Paris : la Découverte. 128 p.
- FORAY, D., & GOENAGA, X. (2013). *The goals of smart specialisation* (Rapport du Joint Research Centre). Luxembourg : Publications Office of the European Union. 18 p.
- FORTUNATO, S. (2010). Community detections in graphs. *Physics and Society*, (486), 75-174. En ligne.
- FORTUNATO, S., & BARTHÉLEMY, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), 36-41. En ligne.
- FRACHISSE, D. (2011, mars 14). *Structures et déterminants des collaborations au sein des programmes cadres de recherche et développement technologique de l'Union Européenne. Une perspective réseau*. (Thèse de Sciences Economiques sous la direction de Pascal Billand et Nadine Massard). Université Jean Monnet de Saint-Etienne, Saint-Etienne, 302 p.
- FRAME, J. D., & CARPENTER, M. P. (1979). International research collaboration. *Social Studies of Science*, 9(4), 481-497.
- FRAME, J. D., & NARIN, F. (1988). The National Self-Preoccupation of American Scientists - An Empirical View. *Research Policy*, 17(4), 203-212. En ligne.
- FRANKHAUSER, P. (1994). *La Fractalité des Structures urbaines*. Paris : Anthropos. 1999. 462 p.
- FREEMAN, A. (2007a). *Defining and measuring metropolitan regions* (Current Issues GLA Economics No. 17). Londres : Greater London Authority. 10 p. En ligne.
- FREEMAN, A. (2007b). *The GLA's interim metro area dataset* (Munich Personal RePEc Archive No. 18130). . 40 p. En ligne.
- FREEMAN, L. C. (1980). Q-analysis and the structure of friendship networks. *International Journal of Man-Machine Studies*, 12(4), 367-378. En ligne.
- FRÉMONT, A. (1986). La géographie sociale, une petite mode qui peut devenir grande. *Espace géographique*, 15(2), 87-89. En ligne.
- FRÉMONT, A. (1990). L'aménagement du territoire universitaire. *Espace géographique*, 193-202. En ligne.
- FRENKEN, K., HARDEMAN, S., & HOEKMAN, J. (2009). Spatial scientometrics: toward a cumulative research program. *Journal of Infometrics*, 3(3), 222-232. En ligne.
- FRENKEN, K., & HOEKMAN, J. (2014). Spatial Scientometrics and Scholarly Impact: A Review of Recent Studies, Tools, and Methods. In Y. DING, R. ROUSSEAU, & D.

- WOLFRAM (DIR.), *Measuring Scholarly Impact* (pp. 127-146). Cham : Springer International Publishing. En ligne.
- FRENKEN, K., HOEKMAN, J., KOK, S., PONDS, R., VAN OORT, F., & VAN VLIET, J. (2009). Death of Distance in Science? A Gravity Approach to Research Collaboration. In A. PYKA & A. SCHARNHORST, *Innovation networks new approaches in modelling and analyzing* (pp. 43-57). Berlin-Heidelberg : Springer-Verlag. En ligne.
- FRENKEN, K., HÖLZL, W., & VOR, F. DE. (2005). The citation impact of research collaborations: the case of European biotechnology and applied microbiology (1988-2002). *Journal of Engineering and Technology Management*, 22(1-2), 9-30. En ligne.
- FRIEDBERG, E. C. (1997). *Correcting the blueprint of life : an historical account of the discovery of DNA repair mechanisms*. Plainview, N.Y. : Cold Spring Harbor Laboratory Press. 210 p.
- FRIEDBERG, E. C. (2005). Reminiscences of a long-time colleague and friend, Philip C. Hanawalt. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 577(1-2), 9-13. En ligne.
- FRIEDBERG, E. C. (2007). A brief history of the DNA repair field. *Cell research*, 18(1), 3-7. En ligne.
- FRIEDMAN, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. New York : Farrar, Straus et Giroux. 496 p.
- GAILLARD, J., & MEYER, J.-B. (1996). Le brain drain revisité : de l'exode au réseau. In *Les sciences hors d'Occident au XXème siècle: Coopérations scientifiques internationales* (Vol. 7, pp. 331-347). Paris : ORSTOM.
- GARCÍA-CARPINTERO, E., GRANADINO, B., & PLAZA, L. M. (2010). The representation of nationalities on the editorial boards of international journals and the promotion of the scientific output of the same countries. *Scientometrics*, 84(3), 799-811. En ligne.
- GARCIA-PARPET, M.-F. (2014). Marché, rationalité et faits sociaux totaux : Pierre Bourdieu et l'économie. *Revue Française de Socio-Économie*, 13(1), 107. En ligne.
- GARCIA-RAMON, M.-D. (2003). Globalization and international geography: the questions of languages and scholarly traditions. *Progress in Human Geography*, 27(1), 1-5. En ligne.
- GARFIELD, E. (1967, décembre 26). English -- an international language for science? *Current Contents*, 19-20.
- GARFIELD, E. (1970a, avril 15). SCI and ASCA promote international contacts. *Current Contents*, (15), 5-6.
- GARFIELD, E. (1970b, septembre 23). Use of Current Contents in Developing Countries. *Current Contents*, (30), 5-6.
- GARFIELD, E. (1972, avril 26). Introducing Morton V. Malin, Ph.D., ISI's vice president for corporate planning & international relations. *Current Contents*, (16), 5-6.

- GARFIELD, E. (1973a, janvier 24). International science requires cultural as well as economic « imperialism ». *Current Contents*, (5-6).
- GARFIELD, E. (1973b, avril 22). The economics and Realpolitik of exponential information growth; or, journal selection ain't easy! *Current Contents*, (36), 5-6.
- GARFIELD, E. (1975, mai 26). Why not stop worrying about Cyrillic and read Russian!, (21), 5-10.
- GARFIELD, E. (1977, avril 16). Le Nouveau Défi Américain. II. *Current Contents*, (16), 5-12.
- GARFIELD, E. (1978). Editorial Statements. *Scientometrics*, 1(1), 5-7.
- GARFIELD, E. (1979, novembre 5). How do we select journals for Current Contents? *Current Contents*, (45), 5-8.
- GARFIELD, E. (1987). Little Science, Big Science...and Beyond Gathers Together the Major Works of Derek de Solla Price. *Current Contents*, (11), 3.
- GARFIELD, E. (2006). Citation indexes for science. A new dimension in documentation through association of ideas. *International Journal of Epidemiology*, 35(5), 1123-1127. En ligne.
- GARFIELD, E., & HAYNE, R. L. (1955). Needed—a national science intelligence & documentation center. In *Symposium on Storage and Retrieval of Scientific Information*. Atlanta, Ga : American Association for the Advancement of Science.
- GARG, K. C. (2003). An overview of cross-national, national, and institutional assessment as reflected in the international journal *Scientometrics*. *Scientometrics*, 56(2), 169-199.
- GARRIGUE, D. (2003). *Recherche européenne : l'urgence du réveil* (Rapport d'information No. 1095). Paris : délégation de l'Assemblée nationale pour l'Union européenne XIIe législature. 144 p. En ligne.
- GASCOIGNE, R. (1992). The historical demography of the scientific community, 1450-1900. *Social Studies of Science*, 22(3), 545-573. En ligne.
- GAUFFRIAUX, M., LARSEN, P., MAYE, I., ROULIN-PERRIARD, A., & INS, M. (2008). Comparisons of results of publication counting using different methods. *Scientometrics*, 77(1), 147-176. En ligne.
- GEORGE, P. (1970). [Recension de *Mark Jefferson, Geographer*, par G. J. Martin]. *Annales de Géographie*, 79(435), 596-598.
- GEORGHIOU, L. (1998). Global cooperation in research. *Research Policy*, 27(6), 611-626. En ligne.
- GHORRA-GOBIN, C. (2012). *Dictionnaire critique de la mondialisation* (2^e éd.). Paris : Armand Colin. 648 p.
- GIBBONS, M., LIMOGES, C., NOWOTNY, H., & SCHWARTZMAN, S. (1994). *The new production of knowledge : The dynamics of science and research in contemporary societies*. London-Thousand Oaks, CA : SAGE Publications. 179 p.
- GILFILLAN, S. C. (1930). Inventiveness by Nation: A Note on Statistical Treatment. *Geographical Review*, 20(2), 301. En ligne.

- GINGRAS, Y. (2002). Les formes spécifiques de l'internationalité du champ scientifique. *Actes de la recherche en sciences sociales*, 2(141), 31-45. En ligne.
- GINGRAS, Y. (2003). Idées d'université. Enseignement, recherche et innovation. *Actes de la recherche en sciences sociales*, 3(148), 3-7. En ligne.
- GINGRAS, Y. (2008). *La fièvre de l'évaluation de la recherche. Du mauvais usage de faux indicateurs* (Note de recherche CIRST). Montréal. 15 p. En ligne.
- GINGRAS, Y. (2013). *Sociologie des sciences*. Paris : Presses universitaires de France. 127 p.
- GINGRAS, Y., & LARIVIÈRE, V. (2011). There are neither « king » nor « crown » in scientometrics: Comments on a supposed « alternative » method of normalization. *Journal of Informetrics*, 5(1), 226-227. En ligne.
- GLÄNZEL, W., DEBACKERE, K., & MEYER, M. (2008). « Triad » or « tetrad »? On global changes in a dynamic world. *Scientometrics*, 74(1), 71-88. En ligne.
- GLÄNZEL, W., & SCHUBERT, A. (2005). Analysing Scientific Networks Through Co-Authorship. In H. F. MOED, W. GLÄNZEL, & U. SCHMOCH (DIR.), *Handbook of Quantitative Science and Technology Research* (pp. 257-276). Dordrecht : Springer Netherlands. En ligne.
- GLÄNZEL, W., THIJS, B., SCHUBERT, A., & DEBACKERE, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics*, 78(1), 165-188. En ligne.
- GLÄSER, J., & LAUDEL, G. (2001). Integrating scientometric indicators into sociological studies: methodical and methodological problems. *Scientometrics*, 52(3), 411-434. En ligne.
- GODECHOT, O. (2012). La formation des relations académiques au sein de l'ehecs. *Histoire & mesure*, 26(2), 223-260.
- GODECHOT, O., & LOUVET, A. (2010). Comment les docteurs deviennent-ils directeurs de thèse ? *Sociologie*, 1, 3-23. En ligne.
- GODIN, B. (2002). Outline for a History of Science Measurement. *Science, Technology, & Human Values*, 27(1), 3-27. En ligne.
- GODIN, B. (2006a). *De l'Eugénisme à la scientométrie*. Montréal. 17 p. En ligne.
- GODIN, B. (2006b). *From Eugenics to Scientometrics: Galton, Cattell and Men of Science* (Document de travail No. 32). Montréal : INRS. 49 p.
- GODIN, B. (2014). *Innovation and Science: When Science Had Nothing to Do with Innovation* (Document de travail No. 16). Montréal : INRS. 31 p. En ligne.
- GONDARD, P. (1991). *Pour une approche géographique de l'innovation : l'innovation est-elle un objet géographique ?* (Essai). Montpellier : ORSTOM. 17 p. En ligne.
- GOTTMANN, J. (1951). *La politique des États et leur géographie*. Paris : Éditions du CTHS. 2007. 261 p.
- GOTTMANN, J. (1957). Megalopolis or the Urbanization of the Northeastern Seaboard. *Economic Geography*, 33(3), 189. En ligne.

- GOTTMANN, J. (1973). The need of an international policy for the sciences. In *Nobel Symposium No. 26* (7-13 p.). Oslo.
- GOULD, P. (1980). Q-analysis, or a language of structure: an introduction for social scientists, geographers and planners. *International Journal of Man-Machine Studies*, 13, 169-199. En ligne.
- GOULD, P., & JOHNSON, J. (1980). National Television Policy Monitoring structural complexity. *Futures*, 12(3), 178-190. En ligne.
- GRANDCLÉMENT, A. (2012, novembre 9). *Géographie des pôles de compétitivité : réseaux et territoires de l'innovation* (Thèse de géographie sous la direction de Sylvie Daviet). Aix-Marseille Université, Aix-en-Provence, 643 p.
- GRANJOU, C., & PEERBAYE, A. (2011). Sciences et collectifs. *Terrains & Travaux*, 1(18), 5-18.
- GRASLAND, C. (2000). Chapitre 1: L'analyse des semis de points. In *Organisation de l'espace (1). Analyse spatiale et modélisation des phénomènes géographiques. Licence de Géographie. Supports de cours*. Paris : Université Paris VII. En ligne.
- GRASLAND, C., & BARON, M. (2006). *Analyse des flux. Représentation cartographique*. (Support de cours). Yaoundé : Université Paris 7 - Yaoundé 1. 46 p.
- GRASLAND, C., & MADELIN, M. (2006). *The modifiable areas unit problem* (ESPON 3.4.3. Final Report). Luxembourg : ESPON. 254 p. En ligne.
- GRATALOUP, C. (1999). Représenter-penser un Monde mondialisé. *Espace géographique*, 28(1), 13-22. En ligne.
- GRATALOUP, C. (2010). *Géohistoire de la mondialisation : Le temps long du monde*. Paris : Armand Colin. 256 p.
- *GREGORY, D. (1979). *Ideology, science, and human geography*. New York : St. Martin's Press. 198 p.
- GRÉSILLON, B. (2014). *Géographie de l'art : ville et création artistique*. Paris : Economica. 256 p.
- GRILICHES, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, 10(1), 92-116. En ligne.
- GROSSETTI, M. (1994). *Université et territoire. Un système local d'enseignement supérieur : Toulouse et Midi-Pyrénées*. Toulouse : Presses universitaires du Mirail. 237 p.
- GROSSETTI, M. (1995). *Science, industrie et territoire*. Toulouse : Presses universitaires du Mirail. 309 p.
- GROSSETTI, M. (1997). *Mémoire pour l'habilitation à diriger des recherches*. Université de Toulouse le Mirail, Toulouse, 128 p.
- GROSSETTI, M. (2005). Where do social relations come from ? : A study of personal networks in the Toulouse area of France. *The Dynamics of Personal Networks*, 27(4), 289-300. En ligne.
- GROSSETTI, M. (2007). Mondes de la science et spécialités artistiques Réflexions sur les formes collectives dans les activités de création. In F. GAUDEZ, *Sociologie des arts, sociologie des sciences* (Vol. 2, pp. 17-27). Paris : L'Harmattan.

- GROSSETTI, M. (2008). *Sociologie des sciences* (Support de cours). Toulouse : Université Toulouse 2. 40 p.
- GROSSETTI, M. (2015). Sur l'émergence des collectifs. In D. DEMAZIÈRE & M. JOUVENET, *Andrew Abbott, sociologue de Chicago, dépassements, ruptures* (sous presse). Paris : Editions de l'EHESS.
- GROSSETTI, M., & BÈS, M.-P. (2002). Proximité spatiale et relations science-industrie : savoirs tacites ou encastrement (Polanyi ou Polanyi)? *Revue d'Économie Régionale & Urbaine*, (5), 777-787.
- GROSSETTI, M., ECKERT, D., GINGRAS, Y., JÉGOU, L., LARIVIÈRE, V., & MILARD, B. (2014). Cities and the geographical deconcentration of scientific activity: A multi-level analysis of publications (1987-2007). *Urban Studies*, 51(10), 2219-2234. En ligne.
- GROSSETTI, M., ECKERT, D., JÉGOU, L., MAISONOBE, M., GINGRAS, Y., & LARIVIÈRE, V. (2013). La diversification des espaces de production du savoir. *CERISCOPE Puissance*. En ligne.
- GROSSETTI, M., ECKERT, D., MAISONOBE, M., & TALLEC, J. (à paraître). Four commonly held beliefs about the geography of scientific activities. In C. CARRINCAZEAX, *Handbook of Regional Innovation and Growth*. Edward Elgar.
- GROSSETTI, M., & FILÂTRE, D. (2003). Que deviendront les nouveaux établissements d'enseignement supérieur et de recherche en France? In M. GROSSETTI & P. LOSEGO, *La territorialisation de l'enseignement supérieur et de la recherche. France, Espagne, Portugal* (pp. 289-318). Paris : L'Harmattan.
- GROSSETTI, M., & GRELON, A. (Dir.). (à paraître). *La construction de l'espace scientifique français*. Paris : CNRS.
- GROSSETTI, M., GRELON, A., BIRCK, F., DÉRÉ, A.-C., DETREZ, C., EMPTOZ, G., ... SPIESSER, M. (1996). *Villes et institutions scientifiques* (Rapport pour le PIR-VILLES). Toulouse : CNRS Midi-Pyrénées. 402 p. En ligne.
- GROSSETTI, M., & LOSEGO, P. (2003). *La territorialisation de l'enseignement supérieur et de la recherche. France, Espagne, Portugal*. Paris : L'Harmattan. 339 p.
- GROSSETTI, M., LOSEGO, P., & MILARD, B. (2009). La territorialisation comme contrepoint à l'internationalisation des activités scientifiques. In P. LAREDO, J.-P. LERESCHE, & K. WEBER, *L'internationalisation des systèmes de recherche en action. Les cas français et suisse* (pp. 281-300). Lausanne : Presses polytechniques et universitaires romandes.
- GROSSETTI, M., & MILARD, B. (2003). Les évolutions du champ scientifique en France à travers les publications et les contrats de recherche. *Actes de la recherche en sciences sociales*, 3(148), 47-56. En ligne.
- GROSSETTI, M., & MILARD, B. (2011). La concentration spatiale des activités scientifiques en question (14 p.). Présenté à 3e colloque international du RESUP (L'enseignement supérieur et la recherche en réformes), Paris : Science Po. En ligne.

- GROSSETTI, M., MILARD, B., & LOSEGO, P. (2002). La territorialisation des activités scientifiques dans le Sud-Ouest européen (France, Espagne, Portugal). *Géographie, Economie, Société*, 4(4), 427-442. En ligne.
- GROSSETTI, M., MILARD, B., & MAISONOBE, M. (à paraître). Une approche socio-historique pour l'étude spatiale des sciences. *Histoire de la recherche contemporaine*.
- GROSSETTI, M., & MOUNIER-KUHN, P.-E. (1995). Les débuts de l'informatique dans les universités : Un moment de la différenciation géographique des pôles scientifiques français. *Revue française de sociologie*, 36(2), 295-324. En ligne.
- GROSSETTI, M., & NGUYEN, D. (2001). La structure spatiale des relations science-industrie en France : l'exemple des contrats entre les entreprises et les laboratoires du CNRS. *Revue d'Economie Régionale et Urbaine*, 2, 311-328. En ligne.
- GROUPE ELEMENTR. (2014). *R et espace. Traitement de l'information géographique* (1^{re} éd.). Paris : Framabook. 244 p. En ligne.
- GUERMOND, Y. (2001). Y a-t-il des territoires innovants? *Réalités Industrielles-Annales des Mines*, 61-67.
- GUÉROIS, M., BRETAGNOLLE, A., MATHIAN, H., & PAVARD, A. (2014). *Functional Urban Areas (FUA) and European harmonization. A feasibility study from the comparison of two approaches: commuting flows and accessibility isochrones* (Technical Report, Espon 2013 Database). Paris : Union Européenne. 35 p.
- GUTIÉRREZ, J., & LÓPEZ-NIEVA, P. (2001). Are international journals of human geography really international? *Progress in Human Geography*, 25(1), 53-69. En ligne.
- HACKING, I. (1980). Is the end in sight for epistemology? *The Journal of Philosophy*, 77(10), 579-588.
- HÄGERSTRAND, T. (1953). *Innovation Diffusion as a Spatial Process [Innovationsförloppet ur korologisk synpunkt]*. (A. PRED & G. HAAG, TRAD.). Chicago : The University of Chicago Press. 1967. 334 p.
- HAGGETT, P., & CHORLEY, R. J. (1967). Models, paradigms and the new geography. In P. HAGGETT & R. J. CHORLEY (DIR.), *Models in geography* (Vol. 19). Londres : Methuen.
- HAGSTROM, W. O. (1965). *The Scientific Community*. Carbondale : Southern Illinois University Press. 1975. 304 p.
- HALL, P., & CASTELLS, M. (1994). *Technopoles of the world: the making of twenty-first-century industrial complexes*. Londres : Routledge. 275 p.
- HANAFI, S. (2011). University systems in the Arab East: Publish globally and perish locally vs publish locally and perish globally. *Current Sociology*, 59(3), 291-309. En ligne.
- HANAWALT, P. (2003). Four decades of DNA repair: from early insights to current perspectives. *Biochimie*, 85(11), 1043-1052. En ligne.
- HANAWALT, P. (2010). Growing up with DNA repair and joining the EMS. *Environmental and Molecular Mutagenesis*, 51(8-9), 890-896. En ligne.
- HARRIS, S. J. (1998). Long-distance corporations, big sciences, and the geography of knowledge. *Configurations*, 6(2), 269-304. En ligne.

- HENNEMANN, S. (2013). Information-rich visualisation of dense geographical networks. *Journal of Maps*, 9(1), 68-75. En ligne.
- HENNEMANN, S., RYBSKI, D., & LIEFNER, I. (2012). The myth of global science collaboration - Collaboration patterns in epistemic communities. *Journal of Informetrics*, 6(2), 217-225. En ligne.
- HENRIQUES, L., & LARÉDO, P. (2013). Policy-making in science policy: The 'OECD model' unveiled. *Research Policy*, 42(3), 801-816. En ligne.
- HENRY, N., & FEKETE, J.-D. (2008). Représentations visuelles alternatives pour les réseaux sociaux. *Réseaux*, 26(152), 59-92. En ligne.
- HÉRAUD, J.-A., MUNIER, F., & RONDE, P. (2006). La densité scientifique et technologique des régions, facteurs de la capacité d'innovation des firmes? In A. RALLET & A. TORRE, *Quelles proximités pour innover?* (pp. 133-149). Paris-Budapest-Kinshasa : l'Harmattan.
- HICKS, D. M., WOUTERS, P., WALTMAN, L., DE RIJCKE, S., & RAFOLS, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- HOEIJMAKERS, J. H. J. (2001). From xeroderma pigmentosum to the biological clock contributions of Dirk Bootsma to human genetics. *Mutation Research*, 485(1), 43-59. En ligne.
- HOEKMAN, J. (2012, octobre 18). *Science in an age of globalisation: The geography of research collaboration and its effect on scientific publishing* (Thèse sous la direction de Koen Frenken et Ron Boschma). Eindhoven University of Technology, Eindhoven, 223 p.
- HOEKMAN, J., FRENKEN, K., & OORT, F. (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721-738. En ligne.
- HOROWITZ, I. (1966). Some aspects of the effects of the regional distribution of scientific talent on regional economic activity. *Management Science*, 13(3), 217-232. En ligne.
- HUANG, M.-H., CHANG, H.-W., & CHEN, D.-Z. (2012). The trend of concentration in scientific research and technological innovation: A reduction of the predominant role of the U.S. in world research. *Journal of Informetrics*, 6(4), 457-468. En ligne.
- HURT, C. D. (1983). A comparison of a bibliometric approach and an historical approach to the identification of important literature. *Information Processing & Management*, 19(3), 151-157. En ligne.
- INHABER, H. (1977). Changes in centralization of science. *Research Policy*, 6, 178-193. En ligne.
- INKSTER, I. (1985). Scientific Enterprise and the Colonial 'Model': Observations on Australian Experience in Historical Context. *Social Studies of Science*, 15(4), 677-704. En ligne.
- JACOB, C. (Dir.). (2007). *Lieux de savoir 1. Espaces et Communautés*. Paris : Albin Michel. 1282 p.

- JAFFE, A. B. (1989). Real Effects of Academic Research. *The American Economic Review*, 79(5), 957-970.
- JAFFE, A. B., TRAJTENBERG, M., & HENDERSON, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3), 577-598. En ligne.
- JALABERT, G., & GRÉGORIS, M. (1987). Turin : de la ville-usine à la technopole. *Annales de Géographie*, 96(538), 680-704. En ligne.
- JÄRV, O., AHAS, R., SALUVEER, E., DERUDDER, B., & WITLOX, F. (2012). Mobile Phones in a Traffic Flow: A Geographical Perspective to Evening Rush Hour Traffic Analysis Using Call Detail Records. *PLoS ONE*, 7(11), e49171. En ligne.
- JEANNIN, P. (2003). De la sociologie de la science aux revues de sociologie. *Bulletin de Méthodologie Sociologique*, 78(1), 5-23. En ligne.
- JEFFERSON, M. (1928). The Civilizing Rails. *Economic Geography*, 4(3), 217-231. En ligne.
- JEFFERSON, M. (1929). The Geographic Distribution of Inventiveness. *Geographical Review*, 19(4), 649-661. En ligne.
- JÉGOU, L. (2014). Toward spatially referenced academic data at global scale: the full geocoding of Wos-Datasets, methods and results. Présenté à 2nd Geography of Innovation International Conference, Utrecht.
- JÉGOU, L., & ECKERT, D. (2008). Quel planisphère de référence pour Google Maps ? *M@ppemonde*, 4(92). En ligne.
- JÉGOU, L., & MAISONOBE, M. (2013). Les collaborations scientifiques mondiales, 1999-2008, une application web de géovisualisation. *M@ppemonde*, 4(112). En ligne.
- JENSEN, P., APOLLONI, A., & ROUQUIER, J.-B. (2011). Copublications scientifiques à distance en Europe. In D. PUMAIN & M.-F. MATTEI, *Données Urbaines* (Vol. 6, pp. 279-287). Paris : Economica/Anthropos.
- JOHNSTON, R. (2006). The politics of changing human geography's agenda: textbooks and the representation of increasing diversity. *Transactions of the Institute of British Geographers, New Series*, 31(3), 286-303. En ligne.
- JOYCE, C. (1981, novembre 5). Cuts could do « irreparable damage » to US research. *New Scientist*, 92(1278), 335.
- JULIEN, P. (2000). Mesurer un univers urbain en expansion. *Économie et Statistique*, 6(336), 3-33. En ligne.
- JULIEN, P. (2002, mars). Onze fonctions pour qualifier les grandes villes. *INSEE Première*, (840), 4.
- JULIEN, P., & PUMAIN, D. (1996). Fonctions stratégiques et images des villes. *Économie et Statistique*, 294(1), 127-135. En ligne.
- JURDANT, B. (1979). Créativité et formation des scientifiques. *Tiers-Monde*, 20(78), 254-255. En ligne.
- KATZ, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, 31(1), 31-43. En ligne.

- KEATING, P., CAMBROSIO, A., & MACKENZIE, M. (1996). Les outils de la discipline : critères, modèles et mesures dans la controverse affinité/avidité en immunologie. In A. E. CLARKE & J. H. FUJIMURA, *La matérialité des sciences : Savoir-faire et instruments dans les sciences de la vie* (pp. 398-450). Le Plessis-Robinson : Synthelabo Groupe.
- KEIM, W. (2010). Pour un modèle centre-périphérie dans les sciences sociales. Aspects problématiques des relations internationales en sciences sociales. *Revue d'anthropologie des connaissances*, 4(3), 570-598. En ligne.
- KENNA, R., & BERCHE, B. (2011). Critical masses for academic research groups and consequences for higher education research policy and management. *Higher Education Management and Policy*, 23(3), 9-29. En ligne.
- KLEICHE-DRAY, M. (Dir.). (à paraître). *Les ancrages nationaux de la science mondiale : Histoire et géographie des activités scientifiques depuis le 18ème siècle*. Paris : Presses de l'IRD.
- KLEIN, J. L., & GUILLAUME, R. (Dir.). (2014). *Vers une nouvelle géographie économique : hommage à Claude Manzagol*. Montréal : Presses de l'Université du Québec. 224 p.
- KNORR-CETINA, K. D. (1982). Scientific Communities or Transepistemic Arenas of Research? A Critique of Quasi-Economic Models of Science. *Social Studies of Science*, 12(1), 101-130. En ligne.
- KOSHLAND, D. (1994). Molecule of the year: the DNA repair enzyme. *Science*, 266(5193), 1925-1925. En ligne.
- KOVALEV, I. (2002). *Geografiâ mirovoj nauki: Gardariki*. Moscou. 156 p.
- KRAEMER, K., & BOHR, V. (2005). The DNA Repair Interest Group: a global village. *DNA Repair*, 4(4), 405-406. En ligne.
- KRUSKAL, J. B. (1964). Multidimensional-scaling by optimizing goodness of fit to a non-metric hypothesis. *Psychometrika*, 29(1), 1-27. En ligne.
- KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional Scaling*. Thousand Oaks, CA : Sage Publications. 93 p.
- KUHN, T. S. (1962). *La structure des révolutions scientifiques*. (L. MEYER, TRAD.) (2^e éd.). Paris : Flammarion. 2008. 284 p.
- KUHN, T. S. (1990). The Road since Structure. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 2, 3-13 p.).
- KUHN, T. S. (2000). *The Road Since Structure: Philosophical Essays, 1970-1993, with an Autobiographical Interview*. (J. CONANT & J. HAUGELAND, DIR.). Chicago : University of Chicago Press. 335 p.
- LABASSE, J. (1982). Lyon, métropole ? *Revue de géographie de Lyon*, 57(2), 145-153. En ligne.
- LABROSSE, I., VENTIMIGLIA, A., CAMPBELL, D., HAUSTEIN, S., CÔTÉ, G., & ARCHAMBAULT, É. (2013). *Scientific Output and Collaboration of European Universities* (Rapport de Science-Metrix pour la Commission Européenne No. EUR26116). Luxembourg : Publications Office of the European Union. 153 p.
- LACOSTE, Y. (1959). *Les Pays sous-développés* (4^e éd.). Paris : Presses universitaires de France. 1963. 127 p.

- LACOUR, C. (2008). Préface. In C. COURLET & B. PECQUEUR, *L'économie territoriale*. Grenoble : Presses universitaires de Grenoble. 142 p.
- LAÏDI, Z. (1996). Le temps mondial. Enchaînements, disjonctions et médiations. *Les Cahiers du CERI*, (14), 4-38.
- LAMBERT, Y. (1995). [Recension de *Globalization. Social Theory and Global Culture*, par R. Robertson]. *Archives des sciences sociales des religions*, (92), 134-136.
- LAMY, J. (2003). La science, le continent ignoré des historiens français? *Cahiers d'histoire. Revue d'histoire critique*, (90-91), 133-151.
- LAMY, J. (2007). *L'Observatoire de Toulouse aux XVIIIe et XIXe siècles : Archéologie d'un espace savant*. Rennes : Presses universitaires de Rennes. 537 p.
- LAMY, J., & SIMOULIN, V. (2015). *In Search of Big Facilities*. à paraître.
- LANCICHINETTI, A., RADICCHI, F., RAMASCO, J. J., & FORTUNATO, S. (2011). Finding statistically significant communities in networks. *PLoS ONE*, 6(4), e18961. En ligne.
- LANDREAU, P. (2012). Gabriel-Jules Delarue (dit Strada). *Bulletin de la société historique et scientifique des Deux-Sèvres*, 4(7).
- LARSEN, P. O., & VON INS, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics*, 84(3), 575-603. En ligne.
- LASSAVE, P., & QUERRIEN, A. (Dir.). (1994). Universités et territoires. *Les Annales de la recherche urbaine*, (62-63), 287 p.
- LATOUR, B. (1996). *Petites leçons de sociologie des sciences*. Seuil. 251 p.
- LATOUR, B., & CALLON, M. (Dir.). (1982). *La science telle qu'elle se fait. Anthologie de la Sociologie des Sciences de langue anglaise*. Paris : PANDORE. 300 p.
- LATOUR, B., & WOOLGAR, S. (1979). *Laboratory life: the social construction of scientific facts*. Beverly Hills : Sage Publications. 271 p.
- LAUDEL, G. (2003). Studying the brain drain: Can bibliometric methods help? *Scientometrics*, 57(2), 215-237.
- LAUDEL, G. (2005). Migration Currents Among the Scientific Elite. *Minerva*, 43(4), 377-395. En ligne.
- LAURENS, P., ZITT, M., & BASSECOULARD, E. (2010). Delineation of the genomics field by hybrid citation-lexical methods: interaction with experts and validation process. *Scientometrics*, 82(3), 647-662. En ligne.
- LAW, J. (1973). The Development of Specialties in Science: the Case of X-ray Protein Crystallography. *Social Studies of Science*, 3(3), 275-303. En ligne.
- LAW, J. (1986). On the methods of long-distance control: vessels, navigation and the Portuguese route to India. *Power, action and belief: A new sociology of knowledge? Sociological Review Monograph*, 32, 234-263.
- LEBEAU, B., & VADELORGE, L. (2012). Enseignement supérieur, recherche et collectivités locales. *Histoire urbaine*, 33(1), 119-131. En ligne.
- LE CHATELIER, H. (1925). *Science et industrie* (Gallica). Paris : Flammarion. 283 p.

- LECLERC, M., & GAGNÉ, J. (1994). International Scientific Cooperation - The continentalization of science. *Scientometrics*, 31(3), 261-292. En ligne.
- *LECLERC, R. (1974). *La décentralisation de la recherche scientifique et technique*. (Thèse de 3e cycle). Université de Rouen, Rouen, 262 p.
- LE DOEUFF, M. (2000). Préface à La nouvelle Atlantide. In F. BACON, M. LLASERA & M. LE DOEUFF (TRAD.), *La nouvelle Atlantide* (2^e éd.). Paris : Flammarion. 177 p.
- LEMAINE, G., MACLEOD, R., MULKAY, M., & WEINGART, P. (1976). *Perspectives on The Emergence of Scientific Disciplines*. Paris-La Haye : Maison des Sciences de l'Homme-Mouton & Co. 275 p.
- LEMERCIER, C. (2005). Analyse de réseaux et histoire. *Revue d'histoire moderne et contemporaine*, 52(2), 88-112.
- LEMERCIER, C., GUZZI-HEEB, S., & BERTRAND, M. (2011). Introduction: où en est l'analyse de réseaux en histoire? *Redes: revista hispana para el análisis de redes sociales*, 21, 12-23.
- LEMERCIER, C., & PICARD, E. (2011). *Quelle approche prosopographique?* (Document de travail). En ligne.
- LEMERCIER, C., & ROSENTAL, P.-A. (2009). *The Structure and Dynamics of Migration Patterns in 19th-century Northern France* (Document de travail). En ligne.
- LÉVY, J. (2008). Six moments de l'invention du monde. *Sciences Humaines*, (11).
- LEVY, R., & BOURBIAUX, O. (2015). La recherche dans les SUVM: un travail d'équipe. In R. LEVY, C. SOLDANO, & P. CUNTIGH, *L'université et ses territoires. Dynamismes des villes moyennes et particularités de sites*. (pp. 97-109). Grenoble : Presses universitaires de Grenoble.
- LEVY, R., HUSSLER, C., & TRIBOULET, P. (2012). The evolution of the French collaborative network of innovation: towards clustering? Présenté à DRUID 2012, Copenhague. En ligne.
- LEVY, R., & JÉGOU, L. (2013). Diversity and location of knowledge production in small cities in France. *City, Culture and Society*, 4(4), 203-216. En ligne.
- LEVY, R., SIBERTIN-BLANC, M., & JÉGOU, L. (2013). La production scientifique universitaire dans les villes françaises petites et moyennes (1980-2009). *M@ppemonde*, 2(110). En ligne.
- LEVY, R., & SOLDANO, C. (2015). Quelles particularités des pratiques d'enseignement supérieur et de recherche dans les Sites universitaires de villes moyennes (SUVM)? In R. LEVY, C. SOLDANO, & P. CUNTIGH, *L'université et ses territoires. Dynamismes des villes moyennes et particularités de sites*. (pp. 13-32). Grenoble : Presses universitaires de Grenoble.
- LEYDESDORFF, L. (1989). The relations between qualitative theory and scientometric methods in science and technology studies. *Scientometrics*, 15(5-6), 333-347. En ligne.

- LEYDESDORFF, L. (1997). Why words and co-words cannot map the development of the sciences. *Journal of the American Society for Information Science*, 48(5), 418-427. En ligne.
- LEYDESDORFF, L., & PERSSON, O. (2010). Mapping the Geography of Science: Distribution Patterns and Networks of Relations among Cities and Institutes. *Journal of the American Society for Information Science and Technology*, 61(8), 1622-1634. En ligne.
- LEYDESDORFF, L., & RAFOLS, I. (2011). Local Emergence and Global Diffusion of Research Technologies: An Exploration of Patterns of Network Formation. *Journal of the American Society for Information Science and Technology*, 5(62), 846-860. En ligne.
- LEYDESDORFF, L., & WAGNER, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2(4), 317-325. En ligne.
- LEYDESDORFF, L., WAGNER, C. S., & BORNMAN, L. (2014). The European Union, China, and the United States in the top-1 % and top-10 % layers of most-frequently cited publications: Competition and collaborations. *Journal of Informetrics*, 8(3), 606-617. En ligne.
- LHOMME, S. (2012). *Introduction au développement sous QGIS* (Support de cours). Paris : Université Paris-Est Créteil. En ligne.
- LIEVROUW, L. A., ROGERS, E. M., LOWE, C. U., & NADEL, E. (1987). Triangulation as a research strategy for identifying invisible colleges among biomedical scientists. *Social Networks*, 9(3), 217-248. En ligne.
- LIMONIER, K. (2012). Analyse géopolitique des enjeux d'une politique de puissance : le cas de la science et de l'innovation en Russie. *Hérodote*, 3(146-147), 193-216. En ligne.
- LIN, C.-S., HUANG, M.-H., & CHEN, D.-Z. (2013). The influences of counting methods on university rankings based on paper count and citation count. *Journal of Informetrics*, 7(3), 611-621. En ligne.
- LIVINGSTONE, D. (2003). *Putting science in its place: Geographies of scientific knowledge*. Chicago : The University of Chicago Press. 234 p.
- LODGE, D. (1975). *Changement de décor*. (Y. COUTURIER & M. COUTURIER, TRAD.) (16^e éd.). Paris : Rivages. 1991. 373 p.
- LODGE, D. (1984). *Un tout petit monde*. (Y. COUTURIER & M. COUTURIER, TRAD.) (9^e éd.). Paris : Rivages. 1992. 428 p.
- LOSEGO, P. (2004). Le travail invisible à l'université : le cas des antennes universitaires. *Sociologie du Travail*, 46(2), 187-204. En ligne.
- LOSEGO, P., & ARVANITIS, R. (2008). La science dans les pays non hégémoniques. *Revue d'anthropologie des connaissances*, 2(3), 334-342.
- LOUAIL, T., LENORMAND, M., ROS, O. G. C., PICORNELL, M., HERRANZ, R., FRIAS-MARTINEZ, E., ... BARTHELEMY, M. (2014). From mobile phone data to the spatial structure of cities. *Nature Scientific reports*, 4. En ligne.
- LOUVEL, S., & LANGE, S. (2010). L'évaluation de la recherche : l'exemple de trois pays européens. *Sciences de la société*, (79), 11-26.

- LUUKKONEN, T., TIJSSEN, R. J. W., PERSSON, O., & SIVERSTSEN, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, 28(1), 15-36. En ligne.
- MACEachREN, A. M., & KRAAK, M.-J. (2001). Research challenges in geovisualization. *Cartography and Geographic Information Science*, 28(1), 3-12. En ligne.
- MACLEOD, R. (1980). On visiting the « Moving Metropolis »: Reflections on the Architecture of Imperial Science. *Historical Records of Australian Science*, 5(3), 1-16. En ligne.
- MACLEOD, R. (1999). Empire and information: Intelligence gathering and social communication in India, 1780-1870. *Minerva*, 37(2), 183-185. En ligne.
- MACROBERTS, M. H., & MACROBERTS, B. R. (1986). Quantitative Measures of Communication in Science: A Study of the Formal Level. *Social Studies of Science*, 16(1), 151-172. En ligne.
- MA, H., FANG, C., PANG, B., & LI, G. (2014). The Effect of Geographical Proximity on Scientific Cooperation among Chinese Cities from 1990 to 2010. *PLoS ONE*, 9(11), e111705. En ligne.
- MAIR, A. (1986). Thomas Kuhn and Understanding Geography. *Progress in Human Geography*, 10(3), 345-369.
- MAISONOBE, M. (2013a). *Analyses dynamiques et longitudinales de réseaux*. Paris : Groupe fmrr. 16 p. En ligne.
- MAISONOBE, M. (2013b). Diffusion et structuration spatiale d'une question de recherche en biologie moléculaire. *M@ppemonde*, 2(110). En ligne.
- MAISONOBE, M. (2015). Emergence d'une spécialité scientifique dans l'espace géographique - La réparation de l'ADN. *Bulletin de Méthodologie Sociologique*, 125, 46-64. En ligne.
- MAISONOBE, M., GIGLIA-MARI, G., & ECKERT, D. (2013). DNA Repair: A changing geography? (1964-2008). *DNA Repair*, 12(7), 466-471. En ligne.
- MAISONOBE, M., GROSSETTI, M., MILARD, B., ECKERT, D., & JÉGOU, L. (à paraître). L'évolution des réseaux de collaborations scientifiques entre villes. Une analyse multi-niveaux des activités scientifiques contemporaines.
- MANIFET, C. (2004). *Gouverner par l'action : le cas des politiques universitaires de La Rochelle, Albi et Rodez* (Thèse de sociologie sous la direction de Daniel Filâtre). Université de Toulouse le Mirail, Toulouse, 577 p.
- MARCHANDISE, S. (2013, décembre 3). *Territoires éphémères. Les réseaux sociaux des étudiants marocains en mobilité internationale* (Thèse de géographie sous la direction d'Henry Bakis). Université Montpellier III-Paul Valéry, Montpellier, 454 p.
- MARCHAND, P. (2014). La Russie et l'espace eurasiatique. *Géoéconomie*, 4(71), 151-166.
- MARTIN-BRELOT, H., GROSSETTI, M., ECKERT, D., GRITSAL, O., & KOVÁCS, Z. (2010). The Spatial Mobility of the 'Creative Class': A European Perspective. *International Journal of Urban and Regional Research*, 34(4), 854-870. En ligne.

- MARTIN, R. (1999). The new « geographical turn » in economics: some critical reflections. *Cambridge Journal of Economics*, 23(1), 65-91. En ligne.
- MARTIN, R., & SUNLEY, P. (2003). Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography*, 3(1), 5-35. En ligne.
- MATTELART, A. (1999). *La communication-monde : histoire des idées et des stratégies*. Paris : la Découverte. 853 p.
- MATTHIESSEN, C. W., & SCHWARZ, A. W. (1999). Scientific Centres in Europe: An Analysis of Research Strength and Patterns of Specialisation Based on Bibliometric Indicators. *Urban Studies*, 36(3), 453-477. En ligne.
- MATTHIESSEN, C. W., SCHWARZ, A. W., & FIND, S. (2002). The Top-level Global Research System, 1997-99: Centres, Networks and Nodality. An Analysis Based on Bibliometric Indicators. *Urban Studies*, 39(5-6), 903-927. En ligne.
- MATTHIESSEN, C. W., SCHWARZ, A. W., & FIND, S. (2010). World Cities of Scientific Knowledge: Systems, Networks and Potential Dynamics. An Analysis Based on Bibliometric Indicators. *Urban Studies*, 9(47), 1879-1897. En ligne.
- MATTHIESSEN, C. W., SCHWARZ, A. W., & SOREN, F. (2002). The Ups and Downs of Global Research Centers. *Science*, 297(5586), 1476-1477. En ligne.
- MELIN, G. (1999). Impact of national size on research collaboration - A comparison between Northern European and American universities. *Scientometrics*, 46(1), 161-170. En ligne.
- MELIN, G. (2000). Pragmatism and self-organization - Research collaboration on the individual level. *Research Policy*, 29(1), 31-40. En ligne.
- MELIN, G., & PERSSON, O. (1998). Hotel cosmopolitan: A bibliometric study of collaboration at some European universities. *Journal of the American Society for Information Science*, 49(1), 43-48. En ligne.
- MERCKLÉ, P. (2004). *Sociologie des réseaux sociaux* (1^{re} éd.). Paris : La Découverte. 121 p.
- MÉRENNE-SCHOUMAKER, B. (1996). *La localisation des services*. Paris : Nathan. 196 p.
- MERTON, R. K. (1938). Science, technology and society in seventeenth century England. *Osiris*, 4(1), 360-632. En ligne.
- MERTON, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. (N. W. STORER, DIR.). Chicago : University of Chicago Press. 605 p.
- MEYER, J.-B., KAPLAN, D., & CHARUM, J. (2001). Nomadisme des scientifiques et nouvelle géopolitique du savoir. *Revue internationale des sciences sociales*, (168), 341-354.
- MEYER, M., & MOLYNEUX-HODGSON, S. (2011). « Communautés épistémiques » : une notion utile pour théoriser les collectifs en sciences? *Terrains & Travaux*, 1(18), 141-154.
- MILARD, B. (2008). L'héritage soviétique dans la nouvelle organisation de la science en Russie. Quels effets sur les pratiques et la valorisation de la recherche? *Revue d'anthropologie des connaissances*, 2(3), 391-411. En ligne.

- MILARD, B. (2010). Les citations scientifiques : des réseaux de références dans des univers de références. L'exemple des articles de chimie. *REDES Revista hispana para el analisis de redes sociales*, 19(4), 69-93.
- MILARD, B. (2011a). *Activités scientifiques, textes et réseaux sociaux. Dynamiques relationnelles à travers les citations, publications et bases de données de la recherche scientifique*. (Habilitation à diriger des recherches). Université Toulouse 2, Toulouse, 222 p.
- MILARD, B. (2011b). Dynamiques relationnelles d'un article scientifique : « Roger et al. (2004) » et ses réseaux. *Terrains & Travaux*, 2(119), 141-160.
- MILARD, B. (2012). Emergence, internationalisation et autonomisation de villes scientifiques secondaires. L'exemple de la chimie à Nantes, Angers et Le Mans. In *Université et Territoires* (pp. 89-104). Rennes : Presses Universitaires de Rennes.
- MILARD, B. (2014). Réseaux et entourages citationnels en chimie : quatre types d'implications dans la recherche. *L'Année sociologique*, 64(1), 15-46. En ligne.
- MILARD, B., & GROSSETTI, M. (2006). L'évolution de la recherche scientifique dans les régions de Russie : déclin ou déconcentration? *M@ppemonde*, 1(81). En ligne.
- MILLER, C. A. (2006). « An Effective Instrument Of Peace »: Scientific Cooperation As An Intrument of U.S. Foreign Policy, 1938-1950. *Osiris*, 21(1), 133-160.
- MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE. (2013). *France Europe 2020 : un agenda stratégique pour la recherche, le transfert et l'innovation*. Paris : Ministère de l'Enseignement Supérieur et de la Recherche. 93 p.
- MÖGLING, D. ALIAS SCHWEIGHARDT CONSTANTIENSEM, T. (1618). *Le Miroir de la Sagesse des Rose-Croix, Speculum Sophicum Rhodo-Stauroticum* (Manifeste Rosicrucien). s. l. 18 p.
- MONTAGNÉ-VILLETTE, S. (2001). Le secondaire est-il soluble dans le tertiaire ? *Annales de Géographie*, 110(617), 22-37. En ligne.
- MORANGE, M. (2003). *Histoire de la biologie moléculaire*. Paris : La Découverte. 369 p.
- MORAVCSIK, M. J. (1966). Some practical suggestions for the improvement of science in developing countries. *Minerva*, 4(3), 381-390. En ligne.
- MORAVCSIK, M. J. (1985). *Strenghtening the coverage of third world science: the bibliographic indicators of the third world's contribution to science*. Philadelphie. 16 p. En ligne.
- MOREAU, M., & TESSON, F. (2011). Modalités et enjeux de l'insertion territoriale d'une université en ville moyenne. Approche à partir de la recherche scientifique à l'université de Pau et des Pays de l'Adour. *Cybergeo: European Journal of Geography*, (570). En ligne.
- MORGAN, K. (2001). The exaggerated death of geography: localised learning, innovation and uneven development. In *The Future of Innovation Studies Conference* (23 p.). Eindhoven : Eindhoven University of Technology. En ligne.
- MORICONI-EBRARD, F. (1991). Les 100 plus grandes villes du monde. *Économie et Statistique*, (254), 7-18. En ligne.
- MORICONI-EBRARD, F. (2000). La loi de métropolisation et la notion de système urbain. In *Colloque Géoponts* (pp. 65-75).

- MORICONI-EBRARD, F., DENIS, E., & MARIUS-GNANOU, K. (2010). «Repenser la géographie économique». Les arrangements du rapport de la Banque Mondiale avec les sciences géographiques urbaines. *Cybergeo: European Journal of Geography*. En ligne.
- MORISSET, B. (1999). Palmarès et classements de villes dans la presse hebdomadaire française : essai sur une géographie du jugement de valeur. *Géographie et cultures*, (29), 3-24. En ligne.
- MORO, B. (2006). *L'organisation territoriale de la recherche publique en Bretagne. Une approche systémique*. (Thèse de géographie sous la direction de Guy Baudelle). Université Rennes 2, Rennes, 389 p.
- MORRIS, S. A., & VAN DER VEER MARTENS, B. (2008). Mapping research specialties. *Annual review of information science and technology*, 42(1), 213-295. En ligne.
- MOUNIER-KUHN, P. (2014). Algol in France: From Universal Project to Embedded Culture. *Annals of the History of Computing, IEEE*, 36(4), 6-25. En ligne.
- MULKAY, M. (1975). Three Models of Scientific Development. *Sociological Review*, 3(23), 509-526. En ligne.
- MULLINS, N. C. (1962). The Development of a Scientific Speciality: The Phage Group and the Origins of Molecular Biology. *Minerva*, 10(1), 51-82. En ligne.
- MUSSELIN, C. (2004). *Les projets européens : coopération ou élitisme?* (Societal No. 44). Paris : Institut de l'entreprise. 3 p. En ligne.
- NAGPAUL, P. S. (2003). Exploring a pseudo-regression model of transnational cooperation in science. *Scientometrics*, 56(3), 403-416.
- NARIN, F., STEVENS, K., & WHITLOW, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21(3), 313-323. En ligne.
- NATARAJAN, A. T. (2007). Mutation Research: The Origin. *Mutation Research*, 635(2-3), 79-80. En ligne.
- NATIONS UNIES. (1998). *Principes et recommandations concernant les recensements de la population et de l'habitat. Première révision* (No. ST/ESA/STAT/SER.M/67/rev.1). New-York : Publication des Nations Unies. 302 p.
- NATIONS UNIES. (2009). *Principes et recommandations concernant les recensements de la population et de l'habitat. Deuxième révision* (No. ST/ESA/STAT/SER.M/67/rev.2). New-York : Publication des Nations Unies. 472 p.
- NATIONS UNIES et EUROSTAT. (2006). *Recommandations de la conférence des statisticiens européens pour les recensements de la population et des logements de 2010* (Rapport de la Commission Economique des Nations Unies pour l'Europe No. ECE/CES/STAT/NONE/2006/4). New-York-Genève : Nations Unies. 229 p.
- NAYLOR, S. (2005). Introduction: historical geographies of science – places, contexts, cartographies. *The British Journal for the History of Science*, 38(1), 1-12. En ligne.
- NEEDHAM, J. (1967). The roles of Europe and China in the evolution of oecumenical science. *Journal of Asian History*, 1(1), 3-32. En ligne.

- NOAILLES, P. (2011). De l'innovation à l'innovateur. Pour une approche structuraliste de l'innovation. *La Revue des Sciences de Gestion*, 1-2(247-248), 13-28. En ligne.
- NOOY, W. DE, MRVAR, A., & BATAGELJ, V. (2005). *Exploratory Social Network Analysis with Pajek*. Cambridge : Cambridge University Press. 334 p.
- Nouvelle économie géographique et géographie : Quel dialogue? (2007). *L'Espace géographique*, 36(3), 193-214.
- NYE, M. J. (1975). The scientific periphery in France: The faculty of sciences at Toulouse (1880-1930). *Minerva*, 13(3), 374-403. En ligne.
- OCDE. (1963). *Manuel de Frascati. Méthode type proposée pour les enquêtes sur la recherche et le développement expérimental*. (6^e éd.). Paris : Les Éditions de l'OCDE. 2002. 292 p. En ligne.
- OCDE. (2012). *Redefining Urban: a new way to measure metropolitan areas*. Paris : Les Éditions de l'OCDE. 148 p. En ligne.
- OCDE. (2013). *Science, technologie et industrie : Tableau de bord de l'OCDE 2013* (Science, technologie et industrie : tableau de bord de l'OCDE). Paris : Les Éditions de l'OCDE. 279 p. En ligne.
- O'KELLY, M. E. (1998). A geographer's analysis of hub-and-spoke networks. *Journal of Transport Geography*, 6(3), 171-186. En ligne.
- OKUBO, Y., MIQUEL, J., FRIGOLETTO, L., & DORÉ, J. (1992). Structure of International Collaboration in Science - Typology of countries through multivariate techniques using a link indicator. *Scientometrics*, 25(2), 321-351. En ligne.
- OLMEDA-GOMEZ, C., PERIANES-RODRIGUEZ, A., ANTONIA OVALLE-PERANDONES, M., GUERRERO-BOTE, V. P., & DE MOYA ANEGON, F. (2009). Visualization of scientific co-authorship in Spanish universities From regionalization to internationalization. *Aslib Proceedings*, 61(1), 83-100. En ligne.
- *OPENSHAW, S. (1984). *The modifiable areal unit problem*. Norwich : Geo Books. 40-41 p.
- OPHIR, A., & SHAPIN, S. (1991). The Place of Knowledge: A Methodological Survey. *Science in Context*, 4(1), 3-21. En ligne.
- OPSAHL, T. (2009, mai 15). *Structure and Evolution of Weighted Networks* (Thèse sous la direction de Pietro Panzarasa et Christian Beck). University of London (Queen Mary College), Londres, 151 p. En ligne.
- OPTHOF, T., & LEYDESDORFF, L. (2010). Caveats for the journal and field normalizations in the CWTS (« Leiden ») evaluations of research performance. *Journal of Informetrics*, 4(3), 423-430. En ligne.
- OPTHOF, T., & LEYDESDORFF, L. (2011). A comment to the paper by Waltman et al., *Scientometrics*, 87, 467-481, 2011. *Scientometrics*, 88(3), 1011-1016. En ligne.
- ORAIN, O. (2006). La géographie comme science. Quand faire école cède le pas au pluralisme. In *Couvrir le monde. Un grand XXe siècle de géographie française* (pp. 90-123). Paris : Ministère des affaires étrangères, Association pour la diffusion de la pensée française (ADPF). En ligne.
- ORAIN, O. (2014). Constructivisme. In *Hypergeo*. GDR Libergéo. En ligne.

- ORFORD, S., HARRIS, R., & DORLING, D. (1999). Geography: Information Visualization in the Social Sciences. A State-of-the-Art Review. *Social Science Computer Review*, 17(3), 289-304. En ligne.
- OST. (2007). *Rapport d'étape sur l'analyse des instructions émises par les établissements à l'attention de leurs chercheurs concernant les publications scientifiques. Projet NormAdresses*. Paris : OST. 23 p. En ligne.
- PAAL, M. (2006). *Maps: harmonised delimitation of the agglomeration areas by morphological criteria, presentation in satellite images as well as in black-and-white maps* (COMET Report Summary). Marburg : CORDIS. En ligne.
- PACI, R., & USAI, S. (2008). Knowledge flows across European regions. *The Annals of Regional Science*, 43(3), 669-690. En ligne.
- PADGETT, J. F. (2011). Marriage and elite structure in Renaissance Florence : 1282-1500. *Redes: revista hispana para el análisis de redes sociales*, 21, 71-97.
- PARADEISE, C., REALE, E., & GOASTELLEC, G. (2009). A Comparative Approach to Higher Education Reforms in Western European Countries. In C. PARADEISE, E. REALE, I. BLEIKLIE, & E. FERLIE (DIR.), *University Governance* (Vol. 25, pp. 197-225). Dordrecht : Springer Netherlands. En ligne.
- PATEL, P., & PAVITT, K. (1994). National Innovation Systems: Why They Are Important, And How They Might Be Measured And Compared. *Economics of Innovation and New Technology*, 3(1), 77-95. En ligne.
- PATY, M. (2001). L'histoire comparative des sciences modernes et le contexte de dépendance. *Convergencia. Revista de Ciencias Sociales (Mexico)*, (24), 11-48.
- PERRIN, L., & SOULARD, O. (2014). *Mission Bay à San-Francisco, nouveau hub de la « santé numérique »* (Note rapide IAU-IdF No. 654). Paris : IAU-IdF. 6 p. En ligne.
- PERSSON, O., GLÄNZEL, W., & DANELL, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421-432. En ligne.
- PESTRE, D. (1995). Pour une histoire sociale et culturelle des sciences. Nouvelles définitions, nouveaux objets, nouvelles pratiques. *Annales. Histoire, Sciences Sociales*, 50(3), 487-522. En ligne.
- PETITJEAN, P. (2001). L'Onu a-t-elle voulu dominer la science mondiale ? Les projets de laboratoires internationaux entre 1946 et 1949. In *Refounding the International Scientific Relations in the Post-War Period: the Exact and Natural Sciences Division of Unesco in its First Years* (1-11 p.). Mexico : XXIe Congrès international d'Histoire des Sciences.
- PETITJEAN, P. (2005). Science and the « Civilizing Mission »: France and the Colonial Enterprise. In B. STUCHTEY, *Science Across the European Empires-1800-1950* (pp. 107-128). Oxford-New York : Oxford University Press-German Historical Institute London. En ligne.

- PEYRON, J. P. (1996). Les délocalisations universitaires et leur importance pour les villes moyennes : l'exemple de Saint-Nazaire et de La Roche-sur-Yon. *Noréis*, 171(1), 557-562. En ligne.
- *PIORE, M. J., & SABEL, C. F. (1984). *The second industrial divide: possibilities for prosperity*. Basic books.
- PIZARRO, N. (1999). Appartenances, places et réseaux de places. *Sociologie et sociétés*, 31(1), 143-161. En ligne.
- PLET, F. (1997). [Recension de *Géographes face au monde. L'Union géographique internationale et les congrès internationaux de géographie.*, par M.-C. Robic, A. M. Briend, & M. Rössler]. *Études rurales*, 166-168.
- PLIEZ, O. (2007). Des jeans chinois dans les rues du Caire, ou les espaces discrets de la mondialisation. *M@ppemonde*, 4(88). En ligne.
- POUX-CHILLÈS, A., & HEINTZ, M. (2012). *Evaluation de la recherche universitaire: Exemple du Research Excellence Framework au Royaume-Uni* (Science & Technologie au Royaume-Uni, journal d'information du Service pour la Science et la Technologie de l'ambassade de France à Londres No. 64). Londres : Ambassade de France à Londres. 31 p.
- PLUME, A., & VAN WEIJEN, D. (2014). *Publish or perish? The rise of the fractional author* (Research Trends. Back to the future: 10 years of Scopus and beyond No. 38). Amsterdam : Elsevier. 16-18 p.
- PLUTNIAK, S., & AGUERA, D. (2013). Reliquats en devenir : une approche dynamique de l'écologie documentaire d'un laboratoire de mécanique des fluides. *Sciences de la Société*, (89), 55-73. En ligne.
- POLANCO, X. (Dir.). (1990). *Naissance et développement de la science-monde: production et reproduction des communautés scientifiques en Europe et en Amérique latine*. Paris : La Découverte. 237 p.
- POLANCO, X. (1995). Les sciences de l'information : bibliométrie, scientométrie, infométrie. *Solaris: Les Presses Universitaires de Rennes*, (2). En ligne.
- POLANYI, M. (1958). *Personal knowledge towards a post-critical philosophy*. Chicago : University of Chicago Press. 2015. 464 p. En ligne.
- POLÈSE, M. (1974). Le secteur tertiaire et le développement économique régional : vers un modèle opérationnel des activités motrices. *L'Actualité économique*, 50(4), 475-490. En ligne.
- PONDS, R. (2008). *Regional Innovation and the Geography of Research Collaboration in Science-based Industries* (Thèse sous la direction de F.G. van Oort et R.A. Boschma). Utrecht University, Utrecht, 165 p.
- PONOMARIOV, B., & TOIVANEN, H. (2014). Knowledge flows and bases in emerging economy innovation systems: Brazilian research 2005-2009. *Research Policy*, 43(3), 588-596. En ligne.

- PONTILLE, D. (2000). *La signature scientifique : espaces d'inscription et mises en ordre* (Thèse de sociologie sous la direction de Jean-Michel Berthelot). Université de Toulouse le Mirail, Toulouse.
- PONTILLE, D. (2003). Authorship Practices and Institutional Contexts in Sociology: Elements for a Comparison of the United States and France. *Science, Technology & Human Values*, 28(2), 217-243. En ligne.
- PONTILLE, D., & TORNAY, D. (2013). La manufacture de l'évaluation scientifique: algorithmes, jeux de données, outils bibliométriques. *Réseaux*, (177), 25-61.
- POSEN, S., & POSEN, J. S. (1969). The Geography of Reprint Requests. *Journal of Medical Education*, 8(44), 648-654. En ligne.
- POWELL, R. C. (2007). Geographies of science: histories, localities, practices, futures. *Progress in Human Geography*, 31(3), 309-329. En ligne.
- POWELL, W., WHITE, D., KOPUT, K., & OWEN-SMITH, J. (2005). Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences. *American Journal of Sociology*, 110(4), 1132-1205. En ligne.
- PRED, A. (1967). Postscript. In T. HÄGERSTRAND, *Innovation Diffusion as a Spatial Process*. Chicago : University of Chicago Press. 334 p.
- *PRICE, D. J. DE S. (1963). *Little science, big science*. New York : Columbia University Press. 119 p.
- PUMAIN, D. (1997). Pour une théorie évolutive des villes. *Espace géographique*, 26(2), 119-134. En ligne.
- PUMAIN, D., & DARDEL, F. (2014). *L'évaluation de la Recherche et de l'Enseignement supérieur*. Paris. 27 p.
- PUMAIN, D., & SAINT-JULIEN, T. (1997). *L'analyse spatiale : localisations dans l'espace* (3^e éd.). Paris : Colin. 2008. 167 p.
- QI, X., DUVAL, R. D., CHRISTENSEN, K., FULLER, E., SPAHIU, A., WU, Q., ... ZHANG, C.-Q. (2013). Terrorist Networks, Network Energy and Node Removal: A New Measure of Centrality Based on Laplacian Energy. *Social Networking*, 2(1), 19-31. En ligne.
- QI, X., FULLER, E., WU, Q., WU, Y., & ZHANG, C.-Q. (2012). Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194, 240-253. En ligne.
- RAFFESTIN, C., & LÉVY, B. (1998). Epistémologie de la géographie humaine. In A. S. BAILLY, *Les concepts de la géographie humaine* (pp. 25-36). Paris : Armand Colin. En ligne.
- RALLET, A., & TORRE, A. (1999). Is geographical proximity necessary in the innovation networks in the era of global economy? *GeoJournal*, 49, 373-380.
- RASS, E., GRABARZ, A., BERTRAND, P., & LOPEZ, B.-S. (2012). Réparation des cassures double-brin de l'ADN, un mécanisme peut en cacher un autre : la ligature d'extrémités non homologues alternative. *Dossier « Radiobiologie » (2e partie), coordonné par Eric Deutsch*, 16(1), 1-10. En ligne.

- RAYNAUD, D. (2010). Why Do Diffusion Data not Fit the Logistic Law? A Note on Network Discreteness, Heterogeneity & Anisotropy. In *From Sociology to Computing in Social Networks: Theory, Foundations and Applications* (pp. 215-230). Vienne : Springer. En ligne.
- REBORA, G., & TURRI, M. (2013). The UK and Italian research assessment exercises face to face. *Research Policy*, 42(9), 1657-1666. En ligne.
- RÉGION PAYS DE LA LOIRE. (2013). *Schéma régional de l'enseignement supérieur, de la recherche et de l'innovation 2014-2020*. Nantes : Direction de l'enseignement supérieur et de la recherche. 49 p. En ligne.
- RETAILLÉ, D. (2012). *Les lieux de la mondialisation*. Paris : Le Cavalier Bleu. 200 p.
- REYNAUD, A. (1976). Les centres de décision des États-Unis. *Espace géographique*, 5(1), 67-72. En ligne.
- REYNAUD, A. (1992). *Une géohistoire. La Chine des Printemps et des Automnes*. Montpellier : RECLUS. 220 p.
- RHEIN, C. (1982). La géographie, discipline scolaire et/ou science sociale ? (1860-1920). *Revue française de sociologie*, 23(2), 223-251.
- ROBIC, M.-C. (2003). La ville, objet ou problème ? La géographie urbaine en France (1890-1960). *Sociétés contemporaines*, 49-50(1), 107. En ligne.
- ROCHEFORT, M. (1957). Méthodes d'étude des réseaux urbains. Intérêt de l'analyse du secteur tertiaire de la population active. *Annales de Géographie*, 66(354), 125-143. En ligne.
- ROGERS, E. M. (1962). *Diffusion of innovations*. New York : Free Press. 1983. 453 p.
- ROLLET, L. (2009). Peut-on faire l'histoire des pôles scientifiques ? *Histoire de l'éducation*, (122), 93-113. En ligne.
- ROMANO, A., & VAN DAMME, S. (2008). Sciences et villes-mondes, XVIe-XVIIIe siècles. Penser les savoirs au large (XVIe-XVIIIe siècles). *Revue d'histoire moderne et contemporaine*, 2(55), 7-18.
- ROMER, P. M. (1986). Increasing returns and long-run growth. *The journal of political economy*, 94(5), 1002-1037. En ligne.
- ROMER, P. M. (1990). Endogenous technological change. *Journal of political Economy*, 98(5), 71-102.
- *RORTY, R. (1979). *Philosophy and the mirror of nature*. Princeton : Princeton University Press. 424 p.
- ROTA, F. S. (2007). Cities as nodes of research networks in Europe. In N. CATTAN, *Cities and Networks in Europe: A critical approach of polycentrism* (pp. 125-138). Montrouge : John Libbey.
- ROTH, C. (2005). Co-evolution in epistemic networks – Reconstructing social complex systems. *Structure and Dynamics: eJournal of Anthropological and Related Science*, 1(3), 1-160.

- ROTH, C. (2008). Réseaux épistémiques : formaliser la cognition distribuée. *Sociologie du travail*, 50(3), 353-371. En ligne.
- ROTTA, R., & NOACK, A. (2011). Multilevel local search algorithms for modularity clustering. *Journal of Experimental Algorithmics*, 16(2.3). En ligne.
- ROYAL SOCIETY. (2011). *Knowledge, networks and nations global scientific collaboration in the 21st century*. London : The Royal Society. 114 p.
- SAINTEVILLE, M. (2009). *Structuration, organisation et territorialisation de l'espace économique par l'activité boursière, ses flux et ses réseaux* (Thèse de géographie sous la direction de Gabriel Dupuy). Université Paris I-Panthéon-Sorbonne, Paris, 818 p.
- SAINT-JULIEN, T. (1990). L'université et l'aménagement du territoire. *Espace géographique*, 19(3), 206-210. En ligne.
- SAINT-JULIEN, T. (1999). *Atlas de France. Volume 10, Services et commerces*. Montpellier-Paris : RECLUS-La Documentation Française. 128 p.
- SAINT-JULIEN, T. (2004). Diffusion Spatiale. In *Hypergéô* (4 p.). En ligne.
- SASSEN, S. (1991). *La ville globale : New york, Londres, Tokyo*. (D.-A. CANAL, TRAD.). Paris : Descartes & Cie. 1996. 533 p.
- SAXENIAN, A. L. (2006). *The New Argonauts: Regional Advantage in a Global Economy*. Cambridge, MA : Harvard University Press. 432 p.
- SCHNEIDER, A., FRIEDL, M. A., & POTERE, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on « urban ecoregions ». *Remote Sensing of Environment*, (114), 1733-1746. En ligne.
- SCHOFER, E., & MEYER, J. W. (2005). *The World-Wide Expansion of Higher Education* (Document de travail No. 32). CDDRL Stanford. 62 p. En ligne.
- SCHOTT, T. (1993). World Science: Globalization of Institutions and Participation. *Science Technology & Human Values*, 18(2), 196-208. En ligne.
- SCHUBERT, A., & BRAUN, T. (1990). International collaboration in the sciences 1981-1985. *Scientometrics*, 19(1), 3-10. En ligne.
- SCHUBERT, A., & BRAUN, T. (1992). Three scientometric etudes on developing countries as a tribute to Michael Moravcsik. *Scientometrics*, 23(1), 3-19. En ligne.
- SCHUBERT, A., & GLÄNZEL, W. (2006). Cross-national preference in co-authorship, references and citations. *Scientometrics*, 69(2), 409-428. En ligne.
- SCOTT, A. J. (2006). Creative cities: conceptual issues and policy questions. *Journal of Urban Affairs*, 28(1), 1-17. En ligne.
- SÉCHET-POISSON, R., & PEYRON, J.-P. (1994). *Les universités de Nantes et de Rennes et leurs antennes : espace imaginé, espace approprié, espace promotionnel*. Rennes II et Nantes. 212 p.
- SETLOW, R. B. (2005). My early days in photobiology with Philip Hanawalt. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 577(1-2), 4-8. En ligne.
- SETLOW, R. B., & SETLOW, J. K. (1961). Soviet Cellular Biophysics 1950-1960. *The Quarterly Review of Biology*, 36(1), 1-49. En ligne.

- SEXTON, J. (2012). State of The World's Science. *Scientific American*, 307(4), 36-40. En ligne.
- SHAPIN, S. (1984). Talking history: Reflections on discourse analysis. *Isis*, 75(1), 125-130. En ligne.
- SHAPIN, S. (1995). Here and Everywhere: Sociology of Scientific Knowledge. *Annual Review of Sociology*, 21(1), 289-321. En ligne.
- SHAPIN, S. (1996). *The Scientific Revolution*. Chicago : University of Chicago Press. 218 p.
- SHAPIN, S. (1998). Placing the view from nowhere: historical and sociological problems in the location of science. *Transactions of the Institute of British Geographers*, 23(1), 5-12. En ligne.
- SHAPIRO, M. A., SO, M., & PARK, H. W. (2009). Quantifying the National Innovation System: Inter-regional Collaboration Networks in South Korea (22 p.). Présenté à 5th International Conference on Webometrics, Informetrics, and Scientometrics & 10th COLLNET meeting, Dalian.
- SHINN, T. (2002). Nouvelle Production du Savoir et Triple Hélice [Tendances du prêt-à-penser les sciences]. *Actes de la recherche en sciences sociales*, 141(1), 21-30. En ligne.
- SHINN, T., & RAGOUET, P. (2005). *Controverses sur la science : pour une sociologie transver-saliste de l'activité scientifique*. Paris : Raisons d'agir. 240 p.
- SHINN, T., VELLARD, D., & WAAST, R. (2010). La division internationale du travail scientifique. *Cahiers de la recherche sur l'éducation et les savoirs*, (9), 7-138.
- SIGRIST, R. (2013). Les communautés savantes européennes à la fin du siècle des Lumières. *M@ppemonde*, 110(2). En ligne.
- SIMMEL, G. (1908a). *Secret et sociétés secrètes*. (P. WATIER & S. MULLER, TRAD.). Paris : Circé. 1996. 119 p.
- SIMMEL, G. (1908b). *Sociologie : études sur les formes de la socialisation*. (L. DEROCHE-GURCEL & S. MULLER, TRAD.). Paris : Presses Universitaires de France. 2010. 754 p.
- SIMOULIN, V. (2012). *Sociologie d'un grand équipement scientifique : Le premier synchrotron de troisième génération*. Paris : ENS éditions. 294 p.
- SMITH, D. A., & WHITE, D. R. (1992). Structure and Dynamics of the Global Economy: Network Analysis of International Trade 1965-1980. *Social Forces*, 70(4), 857-893. En ligne.
- SOPPELSA, J. (1976). Route 128 - Route 495. *Annales de Géographie*, 85(471), 597-617. En ligne.
- SOULARD, O. (2010). *Science Cities : Campus scientifiques et clusters dans les métropoles du XXIe siècle*, Zurich: EtH Science City et Zurich Ouest (No. 1-09-021 et 6-06-010). Paris : IAU-IdF. 51 p. En ligne.
- SOULARD, O., & GUERY, P. (2004). *La Recherche en Ile de France*. Paris : IAU-IdF. 179 p.
- SOSA, S. B., & BRENNAN, J. L. (2014). The UK Research Excellence Framework and the Transformation of Research Production. In C. MUSSELIN & P. N. TEIXEIRA (DIR.),

- Reforming Higher Education* (Vol. 41, pp. 65-80). Dordrecht : Springer Netherlands. En ligne.
- STASZAK, J.-F. (2000). Prophéties autoréalisatrices et géographie. *Espace géographique*, 29(2), 105-119. En ligne.
- STENT, G. S. (1968). That was the molecular biology that was. *Science*, 160(3826), 390-395.
- STERLACCHINI, A. (2008). R&D, higher education and regional growth: Uneven linkages among European regions. *Research Policy*, 37(6-7), 1096-1107. En ligne.
- STIGLITZ, J. E., & WALSH, C. E. (1993). *Principes d'économie moderne. Traduction de la 3ème édition américaine.* (F. MAYER, TRAD.) (2^e éd.). Bruxelles : De Boeck. 2004. 982 p.
- STOCK, W. G. (2011). Informational Cities: Analysis and Construction of Cities in the Knowledge Society. *Journal of the American Society for Information Science and Technology*, 62(5), 963-986. En ligne.
- STOKES, T. D., & HARTLEY, J. A. (1989). Coauthorship, social structure and influence within specialties. *Social Studies of Science*, 19(1), 101-125. En ligne.
- STORER, N. W. (1970). Le caractère international de la science et l'appartenance à une nation. *Revue Internationale Des Sciences Sociales*, 22(1), 80-93.
- STRADA, J. DE. (1899). *Le Paris de l'ère de la science: capitale de l'univers* (Gallica). Paris : Au Temple de la Religion de la Science. 302 p.
- STRASSER, B. J. (2002). Totems de laboratoires, microscopes électroniques et réseaux scientifiques : L'émergence de la biologie moléculaire à Genève (1945-1960) / Laboratory totems, electron microscopes, and scientific networks : The emergence of molecular biology in Geneva (1945-1960). *Revue d'histoire des sciences*, 5-44. En ligne.
- STRIKE, P. (2003). Bacteria and DNA repair - 50 years together. *Microbiology Today*, 30, 8-10.
- SUIRE, R., & VICENTE, J. (2012). Clusters for life or life cycles of clusters : In search of the critical factors of cluster resilience. (34 p.). Présenté à EUROLIO 2012, Saint-Etienne. En ligne.
- SWERTS, E., PUMAIN, D., & DENIS, E. (2012). Subaltern urbanisation in India. *Economic and Political Weekly*, 47(30), 52-62.
- TABARIÉS, M. (2005). *Les apports du GREMI à l'analyse territoriale de l'innovation ou 20 ans de recherche sur les milieux innovateurs* (Cahiers de la MSE). Paris : Maison des Sciences Economiques. 22 p. En ligne.
- TALBOT, D. (2009). L'approche par la proximité : quelques hypothèses et éléments de définitions (25 p.). Présenté à XVIIIème conférence de l'AIMS, Grenoble. En ligne.
- TALLEC, J. (2014, septembre 29). *La construction socio-spatiale de l'innovation en ville moyenne face aux objectifs de compétitivité et d'attractivité des politiques d'aménagements. Les cas des agglomérations d'Albi (Midi-Pyrénées), Alès (Languedoc-Rousillon), Fougères et Quimper*

- (Bretagne). (Thèse de géographie sous la direction de Guy Baudelle et François Taulelle). Université de Toulouse le Mirail, Toulouse.
- TARAMASCO, C., COINTET, J.-P., & ROTH, C. (2010). Academic team formation as evolving hypergraphs. *Scientometrics*, 85(3), 721-740. En ligne.
- TAUVERON, A. (1974). Le « tertiaire supérieur », moteur du développement régional? *Espace géographique*, 3(3), 169-178. En ligne.
- TAYLOR, P. J., HOYLER, M., & EVANS, D. M. (2008). A Geohistorical Study of 'The Rise of Modern Science': Mapping Scientific Practice Through Urban Networks, 1500-1900. *Minerva*, 46(4), 391-410. En ligne.
- TER WAL, A. L. J. (2014). The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure. *Journal of Economic Geography*, 14(3), 589-620. En ligne.
- TER WAL, A. L. J., & BOSCHMA, R. A. (2009). Applying social network analysis in economic geography: framing some key analytic issues. *The Annals of Regional Science*, (43), 739-756. En ligne.
- TESTA, J. (2011). *The Globalization of Web of Science, 2005-2010*. Philadelphie : Thomson Reuters. 26 p. En ligne.
- TESTA, J. (2012). *Le processus de sélection des revues de Thomson Reuters*. Philadelphie : Thomson Reuters. 5 p. En ligne.
- THAKUR, G. S., HUI, P., & HELMY, A. (2013). *GlobalSense: A Framework for Studying Planet-Scale Vehicular Mobility* (Document de travail). Gainesville : University of Florida. 3 p. En ligne.
- THÉRY, H. (2009). Palmarès des universités mondiales, «Shanghai» et les autres. *M@ppemonde*, (96). En ligne.
- TIJSSEN, R. J., WALTMAN, L., & VAN ECK, N. J. (2012). *Research collaboration and the expanding science grid: Measuring globalization processes worldwide* (Document de travail No. 1203.4194). En ligne.
- TÖRNQVIST, G. (1983). Creativity and the renewal of regional life. In *Creativity and context: A seminar report* (Vol. 50, pp. 91-112). Lund : Royal University of Lund.
- TÖRNQVIST, G. (1989). La créativité : une perspective géographique. In M. CHEVALIER (DIR.), *La Géographie de la créativité et de l'innovation : actes du colloque tenu les 9-10 juin 1987 à l'Université de Paris-Sorbonne*. Paris : Université de Paris-Sorbonne, Dép. de géographie.
- TÖRNQVIST, G. (2011). *The geography of creativity*. Northampton, MA : Edward Elgar. 232 p.
- TORRE, A. (2008). On the Role Played by Temporary Geographical Proximity in Knowledge Transmission. *Regional Studies*, 42(6), 869-889. En ligne.
- TRÉPANIER, M., & IPPERSIEL, M.-P. (2003). Hiérarchie de la crédibilité et autonomie de la recherche. L'impensé des analyses des relations universités-entreprises. *Les Actes de la recherche en sciences sociales*, 3(148), 74-82. En ligne.

- TURNBULL, D. (1996). Cartography and Science in Early Modern Europe: Mapping the Construction of Knowledge Spaces. *Imago Mundi*, 48, 5-24. En ligne.
- TURNBULL, D. (1997). Reframing science and other local knowledge traditions. *Futures*, 29(6), 551-562. En ligne.
- UMS RIATE et UMR GÉOGRAPHIE-CITÉS. (2014). ESPON 2013 Database Dictionary of Spatial Units (M4D). En ligne.
- UNESCO. (2010). *UNESCO Science Report 2010. The Current Status of Science around the World*. Paris : UNESCO. 540 p.
- VAN ECK, N. J., & WALTMAN, L. (2009). How to Normalize Co-occurrence Data? An Analysis of Some Well-Known Similarity Measures. *Journal of the American Society for Information Science and Technology*, 60(8), 1635-1651. En ligne.
- VAN ECK, N. J., WALTMAN, L., VAN RAAN, A. F. J., KLAUTZ, R. J. M., & PEUL, W. C. (2013). Citation Analysis May Severely Underestimate the Impact of Clinical Research as Compared to Basic Research. *PLoS ONE*, 8(4). En ligne.
- VAN LEEUWEN, T. N., MOED, H. F., TIJSEN, R. J. W., VISSER, M. S., & VAN RAAN, A. F. J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51(1), 335-346.
- VAN NOORDEN, R. (2010). Cities: Building the best cities for science. *Nature News*, 467(7318), 906-908. En ligne.
- VAN RAAN, A. F. J., VAN LEEUWEN, T. N., VISSER, M. S., VAN ECK, N. J., & WALTMAN, L. (2010). Rivals for the crown: Reply to Opthof and Leydesdorff. *Journal of Informetrics*, 4(3), 431-435. En ligne.
- VARGA, A. (1997). *Regional economic effects of university research: a survey* (Document de travail). Morgantown, WV : West Virginia University, Regional Research Institute. En ligne.
- VASSAL, S. (1969). Les nouveaux ensembles universitaires français : Éléments de géographie urbaine. *Annales de Géographie*, 78(426), 131-157. En ligne.
- VASSAL, S. (1978). [Recension de *La géographie de l'instruction : Peter Meusburger, Entwicklung Stellung und Aufgaben einer Geographie des Bildungswesens. Eine Zwischenbilanz*, par P. Meusburger]. *Annales de Géographie*, 87(481), 348-349.
- VASSAL, S. (1982). [Recension de *Université et ville en République Fédérale Allemande*, par C. Geissler, R. Geipel, & A. Mayr]. *Annales de Géographie*, 91(505), 387-391.
- VASSAL, S. (1988). *L'Europe des universités*. Caen : Editec. 625 p.
- VASSAL, S. (1993). [Recension de *L'École en Europe, Publications Univ. de Caen, 1990, 288 p.*, par U.R.A. C.N.R.S. 915]. *Annales de Géographie*, 102(570), 199.
- VELTZ, P. (1996). *Mondialisation, villes et territoires : l'économie d'archipel*. Paris : Presses universitaires de France. 264 p.
- VENERI, P. (2013). *On City Size Distribution. Evidence from OECD functional urban areas* (Document de travail No. 2013/27). Paris : OCDE. 16 p. En ligne.

- VESSURI, H. (2001). Introduction : la science et ses cultures. *Revue internationale des sciences sociales*, 168(2), 199-206. En ligne.
- VINCK, D. (1999). Les objets intermédiaires dans les réseaux de coopération scientifiques : contribution à la prise en compte des objets dans les dynamiques sociales. *Revue française de sociologie*, 40(2), 385-414. En ligne.
- VISHER, S. S. (1948). Starred Scientists Born or Trained in Indiana. *Proceedings of the Indiana Academy of Science*, 58, 274-287.
- WACKERMANN, G. (1992). *Les Pôles technologiques : une mode ou une nécessité?*. Paris : La documentation française. 142 p.
- WAGNER, C. S. (2005). Six case studies of international collaboration in science. *Scientometrics*, 62(1), 3-26. En ligne.
- WAGNER, C. S. (2008). *The New Invisible College: Science for Development*. Washington, D.C. : Brookings Institution Press. 157 p.
- WAGNER, C. S., & LEYDESDORFF, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, 34(10), 1608-1618. En ligne.
- WAGNER, C. S., & WONG, S. K. (2012). Unseen science? Representation of BRICs in global science. *Scientometrics*, 90(3), 1001-1013. En ligne.
- WALKER, G. C. (2001). Bryn Bridges and mutagenesis: exploring the intellectual space. *Mutation Research*, 485(1), 69-81. En ligne.
- WALLERSTEIN, I. (1982). Le développement du concept de développement. *Sociologie et sociétés*, 14(2), 133-141. En ligne.
- WALLERSTEIN, I. (2004). *Comprendre le monde. Introduction à l'analyse des système-monde*. (CAMILLE HORSEY, TRAD.). La Découverte. 2009. 173 p.
- WALTMAN, L., & ECK, N. J. VAN. (2007). Some comments on the question whether co-occurrence data should be normalized. *Journal of the American Society for Information Science and Technology*, 58(11), 1701-1703. En ligne.
- WALTMAN, L., VAN ECK, N. J., & NOYONS, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Infometrics*, 4(4), 629-635. En ligne.
- WALTMAN, L., VAN ECK, N. J., VAN LEEUWEN, T. N., VISSER, M. S., & VAN RAAN, A. F. J. (2011). Towards a new crown indicator: an empirical analysis. *Scientometrics*, 87(3), 467-481. En ligne.
- WHEELER, P. B. (1982). Revolutions, research programmes and human geography. *The Royal Geographical Society (with the Institute of British Geographers)*, 14(1), 1-6.
- WHITE, H. C., BOORMAN, S. A., & BREIGER, R. L. (1976). Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions. *American Journal of Sociology*, 81(4), 730-780. En ligne.
- WITHERS, C. W. J. (2009). Place and the « Spatial Turn » in Geography and in History. *Journal of the History of Ideas*, 70(4), 637-658. En ligne.

- WOLFF, J.-P. (1996). Villes moyennes et réseaux de communication : l'exemple de Cholet. *Norois*, 695-703. En ligne.
- WOUTERS, P. F. (1998). The signs of science. *Scientometrics*, 41(1-2), 225-241.
- WOUTERS, P. F. (1999). *The citation culture* (Thèse sous la direction de J.J.M. Franse). Université d'Amsterdam, Amsterdam, 278 p. En ligne.
- WRAY, K. B. (2005). Rethinking Scientific Specialization. *Social Studies of Science*, 35(1), 151-164. En ligne.
- WU, X., FU, Q., & ZHANG, H. Y. (2008). How many Chinese journals are included in the newly indexed 700 regional journals on Web of Science? *Journal of Zhejiang University SCIENCE A*, 9(7), 1011-1014. En ligne.
- YATES, F. A. (2001). *The Rosicrucian Enlightenment*. Londres : Routledge. 352 p.
- YI, D. (2007). The coming of reversibility: The discovery of DNA repair between the atomic age and the information age. *Historical Studies in the Physical and Biological Sciences*, 37, 35-72. En ligne.
- ZHOU, P., & LEYDESDORFF, L. (2007). A comparison between the China Scientific and Technical Papers and Citations Database and the Science Citation Index in terms of journal hierarchies and inter-journal citation relations. *Journal of the American Society for Science and Technology*, 58(2), 223-236. En ligne.
- ZIMMERMANN, J.-B. (2008). Le territoire dans l'analyse économique. Proximité géographique et proximité organisée. *Revue française de gestion*, 34(184), 105-118. En ligne.
- ZITT, M., & BASSECOULARD, E. (1999). Internationalization of communication: a view on the evolution of scientific journals. *Scientometrics*, 46(3), 669-685. En ligne.
- ZITT, M., BASSECOULARD, E., & OKUBO, Y. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics*, 47(3), 627-657. En ligne.
- ZITT, M., PERROT, F., & BARRÉ, R. (1998). The transition from « national » to « transnational » model and related measures of countries' performance. *Journal of the American Society for Information Science*, 49(1), 30-42. En ligne.
- ZITT, M., RAMANANA-RAHARY, S., & BASSECOULARD, E. (2003). Correcting glasses help fair comparisons in international science landscape: Country indicators as a function of ISI database delineation. *Scientometrics*, 56(2), 259-282.
- ZUCKERMAN, H. (1988). The sociology of science. In N. J. SMELSER, *Handbook of sociology* (pp. 511-574). Newbury Park, CA : Sage Publications.

Liste des personnes rencontrées

Isabelle This Saint Jean – Vice-présidente chargée de l'enseignement et de la recherche pour l'Ile de France depuis 2010 – Après une Classe Préparatoire en lettres et sciences sociales au lycée Henri IV (promotion de 1984), ITSJ devient docteur en économie. Maître de conférences à Paris 1, elle accepte un poste de professeur à Dunkerque (Université du Littoral), avant de revenir à Paris (Université Paris 13) en 2007 : « l'année de la LRU ». Présidente de l'association « Sauvons La Recherche » à partir de 2008, elle est appelée par Jean-Paul Huchon à la Région Ile-de-France pour occuper le poste qui s'appelait auparavant « vice-présidence de la recherche, de l'enseignement supérieur et de l'innovation ». L'année de sa nomination, « l'innovation passe au développement économique ». Bien que le niveau de la région ne soit pas toujours « l'échelon pertinent », elle considère que « les conditions de vie et d'étude, le dialogue science/société sont des domaines d'action importants pour la région ». Elle précise : « Je ne suis pas pour la régionalité à 100 % : les diplômes et les organismes de recherche sont nationaux, les universités sont françaises ». Ceci-dit : « le cas de l'Ile-de-France est particulier : c'est 40 % du terrain de la recherche ». Opposée à « l'hyper-concentration », elle est pour une « logique de coordination des acteurs », pour la « collégialité » : « il faut inciter et pas punir ! ».

Dirk Bootsma – Directeur du Département de biologie cellulaire et de génétique clinique à l'Université Erasmus de Rotterdam de 1979 à 1990 – Après des études en Biologie à Utrecht, Bootsma soutient une thèse au TNO de Rijswijk sous la direction de J. A. Cohen en 1965. Sa thèse porte sur l'influence des rayons-X sur le cycle cellulaire. Une fois docteur, il devient chef d'équipe dans la nouvelle université de Rotterdam. Aussitôt en poste, il se rend auprès des dermatologues de Rotterdam, qui lui fournissent des cellules de patient XP. Il souhaite découvrir les gènes impliqués dans cette maladie. Ses avancées dans les années 1970 mettent d'autres équipes sur la voie (l'équipe du NCI à Bethesda qui travaillait auparavant sur des cellules bactériennes ainsi que l'équipe de l'Université d'Osaka, la première à cloner le gène ERCC1 à partir d'un cochon d'inde mutant). Au sujet du laboratoire de Rotterdam, il résume : « nous avons été les premiers à utiliser les techniques de la génétique somatique pour analyser et comprendre un syndrome ». Bootsma s'est fait connaître à un large public pour avoir mis en évidence les gènes en cause dans la leucémie. Au cours de sa carrière, il développe des collaborations au niveau européen, il est à l'initiative de la tenue tous les 4 ans de rencontres sur la réparation de l'ADN aux Pays-Bas, de la bibliothèque de cellules de Rotterdam, et enfin de la création du *Dutch Medical Genetics Center*, une école de recherche commune aux universités de Leyde et Rotterdam. À propos de la réparation de l'ADN comme collectif, il déclare : « Il n'y avait pas tellement de compétition dans le domaine de la réparation de l'ADN, d'après moi cela venait de la conférence internationale sur les mécanismes de réparation, tout le monde venait y assister : c'était comme une famille ».

Koos Jaspers – Assistant professeur au Département de génétique de l'Université Erasmus de Rotterdam (Groupe de recherche : *Nucleotide Excision Repair*) – Formé en Chimie à Amsterdam, il est employé au laboratoire de génétique de Rotterdam au milieu des années 1970 (actuel Centre médical Erasmus). Il oriente d'abord ses recherches sur une variante de la maladie XP (XPV), mais est contraint d'abandonner au bout de 4 ans car contrairement aux autres patients XP, les patients XPV présentent tous la même mutation. Puisque les recherches sur le NER « n'avancent pas trop à cette période », il s'intéresse à des cellules de patients sensibles aux rayons-X souffrant d'*Ataxia-Telangiectasia* (AT). Ses travaux de thèse sur AT le mettent en relation avec un pédiatre de Nijmegen confronté à un cas particulier de patient. À partir des cellules de ce patient, il participe à la découverte du « syndrome de Nijmegen ». Pendant son parcours, Jaspers se spécialise dans l'élaboration de tests diagnostiques (prénataux et postnataux) pour les maladies reliées à une déficience du NER (XP, CS, TTD). Il transmet ce savoir-faire à Miria Stefanini, devenue chef d'équipe à Pavie, qui propose elle-aussi des diagnostics. Au sujet des perspectives dans le domaine de la réparation de l'ADN, il explique : « Ce qu'il y avait de plus simple à faire a été fait, il reste maintenant à enquêter sur le rôle de la chromatine. De nouvelles technologies sont disponibles (*transcriptomics*). La maîtrise des nouvelles technologies est déterminante dans ce domaine. ». S'agissant des travaux sur le NER, il précise : « Le groupe NER est devenu très important localement : 1/3 du laboratoire de Rotterdam se consacre au NER. En ce moment, ce qu'il y a à faire sur ce sujet est difficile, il faut une masse critique pour y parvenir. ».

Leon Mullenders – Directeur du laboratoire de toxicologie génétique de l'Université de Leyde depuis 2002 – Il fait des études en Chimie et Biologie au début des années 1970 puis opte pour une thèse sur la réplication de l'ADN à l'Université de Nijmegen. Après avoir travaillé un moment en Chimie Clinique, il décide de « revenir à la vraie recherche » et se lance dans un post-doctorat au laboratoire de mutagenèse de Leyde. Le directeur du laboratoire est alors F. H. Sobels, qui grâce à ses relations à Bruxelles, a sensibilisé les administrateurs de la recherche aux études génétiques sur les radiations et les enjeux environnementaux au niveau national et européen. À partir des années 1980, Mullenders y mobilise les techniques de réplication de l'ADN pour étudier la réparation de l'ADN. En 2000, il succède à P.H. Lohman à la direction du laboratoire. Contrairement au laboratoire de Rotterdam qui fait de la biologie moléculaire (on y clone des gènes, on y isole des protéines), le laboratoire de Leyde est, selon lui, resté spécialisé dans l'interprétation génétique des lésions et des mutations de l'ADN, ce qui fait que « les deux laboratoires sont très complémentaires ». En termes de perspectives, Mullenders confie qu'il « reste beaucoup à faire en biochimie et à mieux comprendre la dynamique des mécanismes de réparation de l'ADN. ». Il précise : « L'étude des dynamiques est un axe de recherche de Rotterdam tandis qu'à Leyde, on cherche à isoler des complexes de protéines. Beaucoup de questions restent ouvertes, comme celle du rôle de la chromatine, dont la compréhension nécessite de nouveaux collaborateurs (Oxford) et de nouvelles compétences (*proteomics*). »

Paul H. Lohman – Directeur du laboratoire de génétique des radiations et mutagenèse de l’Université de Leyde de 1986 à 2002 – Il commence comme étudiant dans le laboratoire de Delft dirigé par W. Berends où il reste de 1958 à 1962 pour se concentrer sur les effets des rayons UV sur l’ADN. Son diplôme en poche, il rejoint le TNO de Rijswijk pour sa thèse. Il s’agit d’une organisation semi-gouvernementale qui, pour des raisons militaires, finance des travaux sur les dangers des radiations et des armes biologiques pour la santé : « l’idée n’est pas de développer des armes, mais d’aider les militaires à se protéger. ». À la mort de J.A. Cohen, directeur du laboratoire médical du TNO, en 1969, et pour des raisons financières, les recherches y deviennent plus appliquées. Avant de prendre la direction de ce laboratoire, qui ferme ses portes dans les années 1980, Lohman fait un post-doctorat à Houston au Texas sur un sujet proche de celui de M.C. Paterson (en post-doctorat à Oak Ridge) : voir si les enzymes capables de reconnaître les lésions d’ADN provoquées par les rayons UV existent aussi chez les humains. Mais, aucune enzyme de ce type n’est trouvée par les deux hommes car, chez l’humain, ce processus est l’œuvre d’un complexe d’enzyme. En 1986, Lohman remplace F. H. Sobels à la direction du laboratoire de mutagenèse de Leyde. Au sujet de l’organisation des recherches sur la réparation de l’ADN, Lohman se satisfait du fait que « les néerlandais sont arrivés à une très sage conclusion : combiner les efforts. Chacun devant choisir son sujet qui permettrait sous un angle ou un autre de mieux comprendre les relations entre la réparation de l’ADN, les lésions d’ADN et le cancer. ».

Harry van Steeg – Chercheur au Centre de protection médicale de l’Institut National pour la Santé Publique et l’Environnement (RIVM) à Bilthoven – Après une thèse en Biochimie à Amsterdam, il fait de la recherche en biologie moléculaire à Utrecht avant de devenir chercheur au RIVM en 1986 où l’enjeu des recherches est plus appliqué. Au début des années 1990, il a l’idée de façonner un animal porteur d’un handicap en matière de réparation de l’ADN, ce pour quoi il va à la rencontre de J.H. Hoeijmakers à Rotterdam (successeur de D. Bootsma). Dès lors, il choisit de travailler sur le groupe XPA sur lequel ne travaille pas l’équipe de Rotterdam mais une équipe d’Osaka. Après que les japonais aient isolé le gène déficient, il produit la première souris mutante porteuse de ce gène. Le résultat est publié en collaboration avec l’équipe japonaise en 1995. Suite à ce travail, il développe des collaborations avec de nombreux partenaires scientifiques dont l’Université de Leyde où il est nommé professeur en 2005. Également intéressé par les mécanismes de vieillissement, il obtient une bourse du NIH pour travailler avec un ancien étudiant de Lohman installé aux États-Unis. Parallèlement, il poursuit des recherches en carcinogenèse avec une équipe de Maastricht. Au sujet de la communauté de la réparation de l’ADN, il note que les collaborations se sont beaucoup diversifiées à l’échelle nationale à l’occasion de la mise au point de modèles murins et à travers le *Dutch Medical Genetics Center*. Il ajoute : « Les souris mutantes ont un rôle dans la recherche appliquée sur le cancer et le vieillissement, recherche qui se développe désormais plus rapidement que les études mécanistiques (fondamentales) en réparation de l’ADN. ».

Hein te Riele – Directeur de la division de réponse au stress cellulaire à l’Institut Néerlandais de Recherche sur le Cancer (NKI) à Amsterdam depuis 2012 – Étudiant en Chimie à Groningen, il fait une thèse sous la direction de G. Venema au sujet des transformations génétiques intra-bactériennes. À la fin des années 1980, il part faire son post-doctorat à l’Institut Jacques Monod à Paris pour étudier la recombinaison de l’ADN. Quand il arrive au NKI d’Amsterdam dans le groupe d’A. Berns, il s’intéresse à la fabrication de souris mutantes. En 1995, il démontre avec M. Radman que l’efficacité de la recombinaison homologue (HR) repose sur l’activation du mécanisme de réparation par mésappariement des bases (MR) et participe à la conception d’une souris mutante déficiente en matière de MR. Dans les années 2000, il entame des recherches sur le syndrome *Fanconi Anemia*. Ses investigations montrent que *Fanconi Anemia* n’est pas lié à un défaut en matière de MR. Au sujet des recherches sur la réparation de l’ADN au Pays-Bas, te Riele commente : « D. Bootsma a joué un rôle important en identifiant les premiers gènes impliqués dans la réparation de l’ADN et il a formé beaucoup de post-doctorants. Désormais, il y a davantage de connexions entre les différentes équipes, et le *Dutch Medical Genetic Center* joue un rôle central en ce sens. De plus, les liens entre les différents mécanismes sont de plus en plus importants ce qui stimule les collaborations. Ceux qui travaillaient sur le NER s’intéressent maintenant au vieillissement, à l’ensemble de la structure cellulaire et à sa dynamique. ».

Ambra Mari – Chercheuse à l’Institut de Pharmacologie et de Biologie Structurale à Toulouse depuis 2009 – Son parcours est décrit en détail dans la seconde section du Chapitre 11. Après un doctorat sous la direction d’A. Sarasin à l’Institut Gustave Roussy (Villejuif), elle effectue un post-doctorat dans le laboratoire de génétique de Rotterdam où elle participe à créer un modèle murin XPB. En 2009, elle est recrutée sur un poste de chercheuse au CNRS en France et constitue l’équipe « Transcription et Réparation de l’ADN » à Toulouse, classée A+ par l’AERES pour le quadriennal 2010-2014. En 2015, elle déménage sur Lyon avec son équipe, dans un nouvel institut rattaché à l’Université Claude Bernard, où elle envisage de reprendre des expériences *in vivo* sur les modèles murins. L’équipe devrait y prendre le nom d’« ExTra » pour « EXcision X TRAnscription ».

Miroslav Radman – Chercheur à l’Institut Jacques Monot à Paris depuis 1998 – Il fait des études en biologie moléculaire à l’Université de Zagreb puis une thèse dirigée par le radiobiologiste M. Errera en Belgique. Entre 1969 et 1973, Radman fait son post-doctorat à Paris auprès de R. Devoret, suivi d’un séjour de recherche à l’Université d’Harvard. Au début des années 1970, il met en évidence le système mutagène SOS, à la suite de quoi il obtient un poste de professeur à l’Université Libre de Bruxelles. Lors d’importantes coupes budgétaires qui paralysent la recherche belge dans les années 1980, il regagne la France et devient directeur de recherche à l’Institut Jacques Monot où il poursuit des recherches sur la réparation de l’ADN par mésappariements des bases (MR), mécanisme qu’il a contribué à découvrir. En 1998, il est nommé professeur de biologie cellulaire à l’Université Paris Descartes et prend la tête de l’unité INSERM « génétique moléculaire, évolutive et médicale ». Au sujet du « clan » de la réparation de l’ADN, il déclare : « Le clan de la réparation manque de notoriété. Il n’y a jamais eu de prix Nobel pour la réparation. ».

CITATIONS

Les citations extraites de l'anglais ont été systématiquement traduites pour être insérées dans le corps du texte. Le texte original est reporté en note de bas de page.

LISTE DES TABLEAUX

Les tableaux empruntés à d'autres auteurs (le cas échéant, la référence est appelée dans le titre) ont été reproduits et traduits en français lorsque le français n'était pas leur langue d'origine.

Tableau 1 – Brillants scientifiques nés en Indiana. Extrait et traduit de l'article « Starred Scientists Born or Trained in Indiana » (Visher, 1948)	30
Tableau 2 – Fiche-résumé du projet « Géoscience ». Tableau repris d'un document de travail réalisé par Mina Kleiche-Dray	112
Tableau 3 – Les entités scientifiques et leur rapport à l'espace. Tableau repris d'un document de travail réalisé par Michel Grossetti et André Grelon (Compte rendu des réunions « Géoscience » des 11 et 12 mai 2011)	116
Tableau 4 – Origine des demandes de réimpressions des articles de Posen et Posen. Extrait et traduit de l'article « The Geography of Reprint Requests » (Posen & Posen, 1969)	136
Tableau 5 – Le SCI versus les estimations de l'UNESCO. Tableau extrait et traduit de l'article « Three scientometric studies on developing countries as a tribute to Michael Moravcsik » (Schubert & Braun, 1992)	139
Tableau 6 – Nombre de revues indexées par l'ISI Thomson Reuters en 2000 et 2007 (moyenne mobile sur 3 ans)	150
Tableau 7 – L'opération « Regional Expansion 2006-2008 ». Source : ISI Thomson Reuters ...	151
Tableau 8 – L'effet du changement de périmètre du Web of Science entre 2000 et 2007 sur la répartition par pays des publications scientifiques totales en 2007 (part mondiale, moyenne mobile sur 3 ans)	153
Tableau 9 – Bilan des trois vagues de géocodage : l'évolution du fichier recette	176
Tableau 10 – Résumé de l'information spatiale issue d'une notice de publication (l'exemple de la Figure 31)	204
Tableau 11 – Les différents moyens de mesurer la participation des entités spatiales contribuant à une même publication (l'exemple de la Figure 31)	205
Tableau 12 – Rang des municipalités versus le rang des agglomérations à partir d'un comptage entier fractionné. Source : Web of Science, moyenne mobile sur 3 ans	208
Tableau 13 – Rang des municipalités en fonction de la méthode de comptage. Source : Web of Science, moyenne mobile sur 3 ans	210
Tableau 14 – Analyse du plus proche voisin pour l'ensemble des villes publiantes	261

Tableau 15 – Analyse du plus proche voisin parmi les villes les plus publiantes.....	262
Tableau 16 – Répartition des 100 villes les plus publiantes en 2000 et 2007 par pays. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	264
Tableau 17 – Évolution par pays des indices de concentration de la production scientifique entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	267
Tableau 18 – La déconcentration intra-régionale des publications scientifiques (1990-2000). Tableau extrait du chapitre « La territorialisation comme contrepoin à l'internationalisation des activités scientifiques » (Grossetti, Losego & Milard, 2009)	302
Tableau 19 – L'évolution du poids des grands centres de production scientifique au sein de leurs espaces nationaux de 1987 à 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	303
Tableau 20 – L'évolution du poids mondial des grands centres de production scientifique en matière de production et de co-production entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	304
Tableau 21 – L'évolution du poids mondial des grands pays de production scientifique en matière de production et de co-production entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	306
Tableau 22 – Nombre d'agglomérations et de pays par publication en 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	310
Tableau 23 – L'évolution de la part nationale des collaborations des grands centres de production scientifique entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	311
Tableau 24 – L'évolution de la part nationale des collaborations des grands pays de production scientifique entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	313
Tableau 25 – La concentration métropolitaine de la science dans les 11 premières nations scientifiques. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	318
Tableau 26 – Évolution de la production des 500 agglomérations les plus publiantes en 2007 par rapport à l'évolution de la production totale entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	320
Tableau 27 – L'évolution de la répartition des publications et des co-publications entre les 500 agglomérations les plus publiantes du monde en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	320
Tableau 28 – Évolution des mesures globales décrivant la connectivité du réseau des 500 agglomérations les plus publiantes du monde en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	323
Tableau 29 – Évolution de la structuration des collaborations scientifiques parmi les 500 premières agglomérations en fonction de leur taille en nombre de collaborations. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	324

Tableau 30 – Les 15 premiers partenaires scientifiques des « villes globales » en nombre de collaborations en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	327
Tableau 31 – L'évolution de la répartition des collaborations scientifiques de l'Union Européenne à 28. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	334
Tableau 32 – L'évolution de la répartition des collaborations scientifiques entre grands espaces de collaboration (2000-2007). Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	337
Tableau 33 – Structure géographique de la production scientifique mondiale par discipline. Source : Web of Science (articles, reviews, letters), moyenne mobile sur 3 ans	341
Tableau 34 – Structure géographique de la production scientifique mondiale par discipline. Suite. Source : Web of Science (articles, reviews, letters), moyenne mobile sur 3 ans	341
Tableau 35 – Profil des agglomérations où sont effectuées des recherches spécialisées sur la réparation de l'ADN (« DNA repair »). Source : Scopus et Web of Science	366
Tableau 36 – Liste des agglomérations urbaines spécialisées dans la réparation de l'ADN entre 1965 et 1975. Source : Scopus	372
Tableau 37 – Liste des institutions spécialisées dans la réparation de l'ADN entre 1965 et 1975. Source : Scopus	373
Tableau 38 – Liste des pionniers de la spécialité ayant signé au moins 10 publications « DNA repair » entre 1965 et 1975. Source : Scopus	373

TABLE DES FIGURES

Les figures empruntées à d'autres auteurs (le cas échéant, la référence est appelée dans le titre) ont été traduites en français lorsque le français n'était pas leur langue d'origine. Le crédit des représentations cartographiques et de réseaux qui participent aux résultats de ce travail est inclus dans leur cadre avec la légende et le titre nécessaires à leur interprétation.

Figure 1 – Le nombre de brevets obtenus à l'étranger en 1925 par les résidents des pays européens pour cent mille habitants. Carte extraite et traduite de l'article « The Geographic Distribution of Inventiveness» (Jefferson, 1929)	28
Figure 2 – Chercheurs ETP (Équivalent temps plein) par région. Carte extraite de l'article « Aménagement du territoire et développement régional : le cas de la recherche scientifique» (Brocard, 1981).....	31
Figure 3 – L'organisation du système de mesure de la science. Schéma extrait et traduit de l'article « Outline for a History of Science Measurement» (Godin, 2002).....	32
Figure 4 – Le modèle linéaire « classique » de l'innovation. Reproduction du schéma de l'article « De l'innovation à l'innovateur : Pour une approche structuraliste de l'innovation» (Noailles, 2011).....	33

Figure 5 – Voisinage et imbrication des terminologies. Schéma extrait de l'article « La définition des industries de haute technologie » (Fache, 1999)	48
Figure 6 – Les dynamiques réflexives des « local buzz » et « global pipeline ». Schéma extrait et traduit de l'article « Buzz-and-Pipeline Dynamics: Towards a Knowledge-Based Multiplier Model of Clusters » (Bathelt, 2007).....	54
Figure 7 – Cartographie des réseaux du pôle PASS à l'échelle régionale. Schéma extrait de la thèse « Géographie des pôles de compétitivité : réseaux et territoires de l'innovation » (Grandclément, 2012, p. 530)	57
Figure 8 – Spécialisations universitaires « régionales » entre changements et reproductions (Baron, 2010)	63
Figure 9 – Carte conçue par Elsevier à partir des données de la base Scopus. Extraite et traduite du rapport « Knowledge, networks and nations » (Royal Society 2011, p. 38).....	104
Figure 10 – Les relations fonctionnelles entre les différentes tâches du programme « Géoscience ». Schéma repris d'un document de travail	113
Figure 11 – Une approche multi-niveaux, quanti-quali et dynamique. Position des objets analysés dans la thèse numérotés par ordre d'apparition	114
Figure 12 – Réseau de collaborations scientifiques des agglomérations russes en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	118
Figure 13 – Du géocodage à l'analyse spatiale des données bibliographiques. Schéma extrait d'une présentation de Laurent Jégou, Montréal, octobre 2011.....	123
Figure 14 – Croissance relative des publications, des auteurs, des références et des citations (tous domaines confondus). Extraite et traduite de l'article « Inflationary bibliometric values : The role of scientific collaboration and the need for relative indicators in evaluative studies » (Persson, Glänzel, & Danell, 2004).	127
Figure 15 – La couverture du Web of Science en 2008 : comptage du nombre de revues indexées par an. Source : ISI Thomson Reuters	137
Figure 16 – La couverture des grands index du Web of Science en 2008 : comptage du nombre de revues indexées par an et par index. Source : ISI Thomson Reuters	138
Figure 17 – Évolution des langues dans le SCI, 1980-1998. Figure extraite et traduite de l'article « Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance » (van Leeuwen et al., 2001) .	143
Figure 18 – Pôle d'excellence universitaire, répartition géographique des établissements les plus influents. Figure extraite du rapport de l'OCDE « Science, Technologie et Industrie : Tableau de bord de l'OCDE 2013 » (OCDE, 2013, p. 54)	164
Figure 19 – La chaîne de structuration des données géographiques. Schéma repris et traduit du proceeding « Towards spatially referenced academic data at global scale : the full geocoding of Wos-Datasets, methods and results » (Jégou, 2014).....	173
Figure 20 – L'outil d'aide à la vérification du géocodage. Capture d'écran témoignant d'une erreur de géolocalisation : un point « Guadeloupe » dans l'Isère, France	174

Figure 21 – Extrait du fichier recette n°13 permettant d'associer à un triplet source (ville-province-pays) un triplet destination et un couple de coordonnées géographiques	176
Figure 22 – Publications dans le Web of Science, localisations et zones urbaines, variété des référencements. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)	178
Figure 23 – Les trois limites successives de l'agglomération parisienne en 1960. Schéma extrait de l'article « Délimitation de l'agglomération parisienne » (Bastie & Brichler, 1960).....	182
Figure 24 – Les étapes d'une définition harmonisée de l'agglomération. Schéma extrait de l'article « Les 100 plus grandes villes du monde » (Moriconi-Ebrard, 1991)	183
Figure 25 – Les différents processus pour implémenter les modèles conceptuels d'aires urbaines fonctionnelles. Schéma traduit et adapté du rapport technique « Functional Urban Areas (FUA) and European harmonization » (Guérois, Bretagnolle, Mathian, & Pavard, 2014) ...	187
Figure 26 – La détection et la délimitation automatisée de « taches » urbaines au regard de la répartition spatiale des publications scientifiques : l'exemple de Toulouse par Laurent Jégou	193
Figure 27 – Critères de délimitation des zones urbaines, variation des résultats. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)	195
Figure 28 – L'outil d'aide à la vérification du géocodage et d'assistance à la délimitation des « agglomérations scientifiques ».....	196
Figure 29 – Publications dans le Web of Science et hiérarchie mondiale. Figure extraite de l'article « Les villes et la science : apports de la spatialisation des données bibliométriques mondiales » (Eckert et al., 2011)	197
Figure 30 – Répartition des agglomérations scientifiques sur la terre habitée. Source : WoS et CIESIN.....	200
Figure 31 – Notice de publication signée par 16 chercheurs répartis entre Toulouse et Paris. Source : Web of Science	203
Figure 32 – Carte de la production scientifique néerlandaise en 2007 par municipalité (309 municipalités publiantes). Source : Web of Science, moyenne mobile sur 3 ans	207
Figure 33 – La délimitation des « agglomérations scientifiques » dans la Randstad. Capture d'écran de l'outil de géovisualisation	208
Figure 34 – Le monoligne et les effets sur l'indexation des traitements automatiques de l'ISI Thomson Reuters. Extrait du rapport d'étape du groupe de travail « Analyse des Instructions » pour le projet Norm-Adresse (OST, 2007).....	214
Figure 35 – La perte d'information dans la transformation d'un réseau two-mode en réseau one-mode. Schéma extrait de la thèse d'Antoine Grandclément (Grandclément, 2012, p. 200-207).....	220
Figure 36 – Le passage d'un graphe bipartite à un treillis de Galois. Schéma extrait de l'article « Réseaux épistémiques : formaliser la cognition distribuée » (Roth, 2008)	221

Figure 37 – L’effet sur la structure du réseau de co-publications (1990-2007) de l’agrégation des données à différents niveaux géographiques. Source : Web of Science	225
Figure 38 – Les « amitiés » Facebook par Paul Butler.....	237
Figure 39 – Les collaborations scientifiques dans le monde par Olivier Beauchesne, 2011	238
Figure 40 – Une géovisualisation récente des données de production scientifique, via Google Earth avec l’aimable autorisation de l’auteur (© L. Leydesdorff) (Eckert et al., 2013)	241
Figure 41 – Répartition géographique des publications scientifiques dans le monde en 2007. Source : Web of Science, moyenne mobile sur 3 ans	243
Figure 42 – Composante principale du réseau mondial des collaborations scientifiques en 2007. Source : Web of Science, moyenne mobile sur 3 ans	243
Figure 43 – Présentation générale de la recherche et histogrammes	246
Figure 44 – Matrice des collaborations scientifiques entre agglomérations en 2007	246
Figure 45 – Diagramme de chord des collaborations scientifiques entre pays en 2007	247
Figure 46 – Carte de flux des collaborations scientifiques entre agglomérations en 2007	248
Figure 47 – Résultat du partitionnement de la matrice de collaboration scientifique entre agglomérations en 2007, obtenu en appliquant la méthode de détection de communautés « VoS Clustering », une variante de la méthode de Louvain	249
Figure 48 – Graphe des collaborations scientifiques entre agglomérations en 2007.....	250
Figure 49 – Extrait de la table indiquant le nombre de publications par agglomération par année. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	254
Figure 50 – Extrait de la matrice de collaborations scientifiques entre agglomérations à l’année 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	255
Figure 51 – Forme de la distribution d’après la valeur de l’indice de dispersion (R_n). Figure extraite et traduite du support de cours de Claude Grasland (Grasland, 2000).....	260
Figure 52 – Déplacement du centre de gravité de la production scientifique mondiale. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	263
Figure 53 – La distribution de la production scientifique des 100 premières villes en 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	265
Figure 54 – Extrait de la liste des liens de collaborations scientifiques internationales en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	272
Figure 55 – La mise en forme d’un graphe sous Pajek : deux types de fichiers sources (net et mat) décrivant le même graphe. Figure reprise d’une présentation de Andrej Mrvar	273
Figure 56 – Vues locales et globales d’un réseau. Figure reprise d’une présentation de Andrej Mrvar portant sur le logiciel Pajek	277
Figure 57 – L’interface de visualisation VoSviewer. L’exemple des collaborations scientifiques néerlandaises. Source : Web of Science, 2006-2008.....	282

Figure 58 – Cartes de la production scientifique mondiale et son évolution entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	307
Figure 59 – Carte de la production scientifique européenne et son évolution entre 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	308
Figure 60 – Corrélation entre l'évolution du poids dans la production scientifique et la part nationale des collaborations des 30 premiers centres scientifiques mondiaux. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	314
Figure 61 – Corrélation entre l'évolution du degré de concentration de la production nationale par agglomération (indice de Gini) et la part nationale des collaborations des 36 premiers pays du monde. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans ..	315
Figure 62 – Quelles sont parmi les 500 premières agglomérations scientifiques celles qui participent à la croissance de la production des publications entre 2000 et 2007 ? Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	321
Figure 63 – Quelles sont parmi les 500 premières agglomérations scientifiques celles qui participent à la croissance des collaborations entre 2000 et 2007 ? Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	322
Figures 64 et 65 – Le haut de la hiérarchie du réseau des 500 premières agglomérations en 2000 et 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	325
Figure 66 – Dendrogramme issu d'une détection de communautés de type Louvain appliquée pas à pas au réseau interurbain des collaborations scientifiques mondiales entre les 500 premières agglomérations en 2007. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	330
Figure 67 – Positionnement multiscalaire des villes européennes et des villes sous influence européenne d'après l'intensité de leurs collaborations scientifiques. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans.....	332
Figure 68 – Positionnement multiscalaire des pays dans le réseau des collaborations scientifiques mondiales en 2000. Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	336
Figure 69 – Carte de l'évolution des collaborations scientifiques entre grands espaces de collaborations mondiaux (2000-2007). Source : SCI Expanded (articles, reviews, letters), moyenne mobile sur 3 ans	338
Figure 70 – Nombre de publications « DNA repair » par an et par source bibliographique	354
Figure 71 – Répartition des 1675 publications avec le mot-clef « DNA Repair » (1960-1975). Source : Scopus	355
Figure 72 – La répartition géographique des publications « DNA repair » en 2006-2008. Source : Web of Science.....	357
Figure 73 – Lieux d'où sont publiés des documents avec le mot-clef « DNA repair » en 2006, 2007 et 2008. Source : Scopus et Web of Science.....	357
Figure 74 – L'évolution de la part de publications avec le mot-clef « DNA repair ». Source : Scopus et Web of Science	359

Figure 75 – Zoom sur l'évolution de la part de publications avec le mot-clef « DNA repair ». Source : Scopus et Web of Science	360
Figure 76 – Différentes lésions et mécanismes chargés de leur réparation. Figure extraite de l'article « Réparation des cassures double-brin de l'ADN, un mécanisme peut en cacher un autre : la ligature d'extrémités non homologues alternative. » (Rass, Grabarz, Bertrand, & Lopez, 2012)	364
Figure 77 – Graphe bipartite auteurs-institutions synthétisant les parcours institutionnels suivis par 35 auteurs pionniers dans l'utilisation du mot-clef « DNA repair » entre 1960 et 1975. Extrait de l'article « Émergence d'une spécialité scientifique dans l'espace géographique - La réparation de l'ADN » (Maisonobe, 2015).....	375
Figure 78 – Les trajectoires internationales des pionniers de la réparation de l'ADN (1960-1975). Source : Scopus et Web of Science	379
Figure 79 – Connexions internationales d'après les co-signatures enregistrées dans le corpus « DNA repair » entre 1965 et 1975, extrait de la base Scopus	387
Figure 80 – Rythme de réception et d'adoption du mot-clef « DNA repair » (1958-1980)	389
Figure 81 – Réseau des références de l'article « DNA repair » de B.A. Bridges (paru en 1966)	391
Figure 82 – Carte indiquant la destination des cargaisons de radioisotopes de la Commission de l'Énergie Atomique des États-Unis lors des 22 premiers mois du programme. Extraite et traduite de l'article « Tracing the politics of changing postwar research practices: the export of 'American' radioisotopes to European biologists » (Creager, 2002).....	394
Figure 83 – Notice de publication de l'article démontrant le lien entre réparation et transcription de l'ADN. Source : Web of Science	400
Figure 84 – Corpus de publications extraites du Web of Science à l'aide de mots-clefs sur le lien entre réparation et transcription de l'ADN : sélection de deux périodes d'étude	401
Figure 85 – La répartition des agglomérations publiantes selon l'année de leur première publication relative à la question de recherche. Source : Web of Science.....	402
Figure 86 – Nombre annuel de nouvelles agglomérations publiantes sur la période étudiée. Source : Web of Science	404
Figure 87 – Carte de l'évolution de la production scientifique sur la question de recherche (1990- 2007). Source : Web of Science	405
Figure 88 – L'évolution de la structure du réseau de co-publications dans le temps : 1991-1999 : Leyde-Rotterdam au centre (liens supérieurs à 2). Source : Web of Science.....	409
Figure 89 – L'évolution de la structure du réseau de co-publications dans le temps : 2000-2007 : Une organisation polycentrique (liens supérieurs à 2). Source : Web of Science.....	410
Figure 90 – Réseau de références qualifié de l'article scientifique d'AM commenté en entretien	424
Figure 91 – Réseau de co-publication des 360 publications citant au moins 3 références communes à l'article scientifique commenté en entretien. Source : Web of Science	428
Figure 92 – Schéma récapitulant la chaîne de traitement spatial des données bibliographiques .	435

TABLE DES ENCADRÉS

Encadré 1 - La dualité Rostow-Vernon. Schéma extrait de l'Introduction à l'ouvrage collectif <i>Les Régions qui gagnent</i> (Benko, p. 20-21 in Benko & Lipietz, 1992)	41
Encadré 2 - La loi de Lotka. Encadré extrait de l'article « Diffusion et structuration spatiale d'une question de recherche en biologie moléculaire » (Maisonobe, 2013b)	228
Encadré 3 - Distinguer concentration statistique et concentration spatiale. Encadré extrait de l'article « Mesurer un univers urbain en expansion » (Julien, 2000).....	258

Résumé

En géographie et en aménagement, la recherche scientifique est généralement abordée comme un levier de développement économique ou comme un service associé aux établissements de formation supérieure. Cette thèse propose d'envisager la géographie des activités scientifiques à travers leur dimension productive (les publications), sans toutefois réduire l'analyse des logiques spatiales caractéristiques de cette production à des logiques économiques. Une méthode originale de localisation des publications contenues dans les index de citations est proposée qui présente trois particularités : retenir le niveau local comme niveau élémentaire d'analyse pour étudier la répartition et l'organisation des recherches à l'échelle mondiale, définir une population d'agglomérations urbaines comparables auxquelles associer les données de production, s'appuyer sur le nombre de co-signatures entre auteurs pour en déduire des réseaux de collaboration entre lieux. Les résultats obtenus portent sur le contenu de la production scientifique recensée dans le *Science Citation Index Expanded (Web of Science)* entre 1999 et 2008. Ils montrent un mouvement de diffusion de l'activité à plusieurs niveaux, atténuant le monopole des espaces autrefois hégémoniques et sur-représentés dans cet ensemble de références mondialement visibles. Une étude de cas est réalisée sur un domaine de recherche en biologie moléculaire : la réparation de l'ADN. Considérant le rôle des trajectoires individuelles, elle montre qu'une part importante des lieux d'émergence de la spécialité continue de jouer un rôle dans son développement, y compris dans le cadre d'une question de recherche associée au domaine de la transcription de l'ADN.

Mots-clefs : géographie des sciences, production scientifique, réseaux de lieux, agglomérations urbaines, dynamiques de recherche

Abstract

Towards a geography of scientific activities and collectives in the world.

From the growth of the contemporary production system to the dynamics of a specialty: DNA repair.

In geography, scientific research is often conceived as a lever for economic development or a service activity related to higher education. This thesis proposes to consider the geography of scientific activities through its productive dimension (publications). To localize scientific publications retrieved in citations index, an original method is designed which relies on three principles: taking the local level as an elementary level of analysis to study the repartition and organization of research activity at the world scale, defining a set of urban areas built on homogeneous criteria enabling comparison, taking into account co-authorship data to deduce networks of scientific collaborations between places. The main results are derived from the scientific production retrieved in the *Science Citation Index Expanded (Web of Science)* between 1999 and 2008. They show a trend toward the spatial diffusion of production activity at several levels, mitigating the monopoly of hegemonic and over-represented areas in this set of world widely visible scientific references. A case study is focusing on a research field in molecular biology: DNA Repair. Considering the role of individual trajectories, it shows that most pioneering places are still playing a role in the development of the scientific specialty, notably in the frame of a problem area related to the field of DNA Transcription.

Keywords: geography of science, scientific production, network of places, urban areas, research dynamics